

# SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
 Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_  
 Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

**If more than one search is submitted, please prioritize searches in order of need.**

\*\*\*\*\*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

*\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

**BEST AVAILABLE COPY**

Jan Delaval  
 Reference Librarian  
 Biotechnology & Chemical Library  
 CM1 1E07 - 703-308-4498  
 jan.delaval@uspto.gov

\*\*\*\*\*  
**STAFF USE ONLY**

	Type of Search	Vendors and cost where applicable
Searcher: <u>Jan</u>	NA Sequence (#) _____	STN _____
Searcher Phone #: <u>4498</u>	AA Sequence (#) <input checked="" type="checkbox"/>	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>5/1/03</u>	Bibliographic _____	Dr.Link _____
Date Completed: <u>5/1/03</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems <input checked="" type="checkbox"/>
Clerical Prep Time: <u>20</u>	Patent Family _____	WWW/Internet _____
Online Time: <u>15</u>	Other _____	Other (specify) _____

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:31:42 ; Search time 27.2873 Seconds  
(without alignments)  
1602.986 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 2375  
Sequence: 1 MSSLGASFWQIKFDDIQFFE.....GDDDDDDGEEEDNDMNS 455

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

al number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73:\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	601	25.3	954	1 S68178	mixed-lineage prot
2	593	25.0	394	2 J00229	mixed-lineage prot
3	581.5	24.5	847	1 A53800	mixed-lineage prot
4	565.5	23.8	668	2 JC2363	protein kinase (EC
5	565.5	23.8	888	2 A55318	serine/threonine p
6	560.5	23.6	406	2 T52626	probable mitogen-a
7	560.5	23.6	888	2 JC5399	dual leucine zipper
8	545.5	23.0	328	2 T16747	hypothetical prote
9	529	22.3	886	2 T48544	MAP3K delta-1 prot
10	510.5	21.5	1030	2 F96763	hypothetical prote
11	505	21.3	848	2 B87950	protein F33E2.2 (I
12	505	21.3	855	2 T20082	hypothetical prote
13	495	20.8	982	2 T06576	probable protein k
14	488	20.5	821	2 T48400	serine/threonine-p
15	483	20.3	579	2 JC5955	transforming growt
16	480	20.2	829	2 T07406	probable protein k
17	478.5	20.1	1015	2 T07406	probable serine/th
18	475.5	19.9	412	2 T10671	protein kinase hom
19	471.5	19.9	567	2 JC5957	transforming growt
20	471.5	19.9	606	2 JC5956	transforming growt
21	470	19.8	736	2 T05137	protein kinase hom
22	468	19.7	390	2 T01451	protein kinase hom
23	466	19.6	462	2 G29851	protein kinase 6 (
24	460	19.4	963	2 T09911	probable serine/th
25	454.5	19.1	553	2 T04683	hypothetical prote
26	447.5	18.8	407	2 G84635	probable protein k
27	445.5	18.8	546	2 D84555	probable protein k
28	441	18.6	475	2 T12955	probable protein k
29	434	18.3	981	1 F0MVM	gag-abl polyprotei

30	424	17.9	1257	2 T00486	serine/threonine-s
31	421.5	17.7	738	2 F96701	hypothetical prote
32	410	17.3	1130	1 TVH0A	protein-tyrosine k
33	407	17.1	842	2 T32258	hypothetical prote
34	405.5	17.1	697	2 A26132	gag-abl-pol polypr
35	405	17.1	1123	2 A39962	kinase-related tra
36	405	17.1	1146	2 B35962	protein-tyrosine k
37	405	17.0	1182	2 A35962	protein-tyrosine k
38	404	17.0	988	2 F86316	protein T10022.13
39	403.5	17.0	1147	2 F86297	hypothetical prote
40	402	16.9	1520	1 TVPFA	protein-tyrosine k
41	401.5	16.9	1584	2 T18276	protein-tyrosine k
42	400.5	16.9	1094	2 S49313	protein kinase - s
43	400	16.8	1248	2 B96827	hypothetical prote
44	398.5	16.8	1171	2 T12956	hypothetical prote
45	396.5	16.7	545	2 T05675	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

S68178  
Mixed-lineage protein kinase 2 (EC 2.7.1.-) - human

C/Species: Homo sapiens (man)

C/Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 11-Jun-1999

C/Accession: S68178; I38044; S32468

R/Dorow, D.S.; Devereux, L.; Tu, G.F.; Price, G.; Nicholl, J.K.; Sutherland, G.R.; Stamp

Eur. J. Biochem. 234, 492-500, 1995

A/Title: Complete nucleotide sequence, expression, and chromosomal localization of huma

A/Reference number: S68178; MUID:96128179; PMID:8536694

A/Accession: S68178

A/Molecule type: mRNA

A/Residues: 1-954 <DOR>

A/Cross-references: EMBL:X00846; NID:9971419; PIDN:CA62351.1; PID:9971420

R/Katoh, M.; Hirai, M.; Sugimura, T.; Terada, M.

Oncogene 10, 1447-1451, 1995

A/Title: Cloning and characterization of MST, a novel (putative) serine/threonine kinase

A/Reference number: I38044; MUID:95249256; PMID:7731697

A/Accession: I38044

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-461, 'A', 'V', '465-470, 'S', '472-806, 'R', '808-817, 'A', '819-954 <RES>

A/Cross-references: EMBL:Z48615; NID:9758592; PIDN:CA88531.1; PID:9758593

R/Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.

Eur. J. Biochem. 213, 701-710, 1993

A/Title: Identification of a new family of human epithelial protein kinases containing

A/Reference number: S32467; MUID:93238756; PMID:8477742

A/Accession: S32468

A/Molecule type: mRNA

A/Residues: 244-464, 'AQAARRQPHQPALML', <DO2>

C/Genetics:

A/Gene: GDB:MK2; GDB:MST

A/Cross-references: GDB:362654; GDB:624810; OMIM:600137

A/Map position: 19q13.1-19q13.2

C/Superfamily: mixed-lineage protein kinase 2; protein kinase homology; SH3 homology

C/Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein

F:23-76/Domain: SH3 homology <SH3>

F:96-364/Domain: protein kinase homology <KIN>

F:104-112/Region: protein kinase ATP-binding motif

F:384-405/Region: leucine zipper motif

F:419-440/Region: leucine zipper motif

F:449-463/Region: basic

F:125,145,222,224/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match

Best Local Similarity 25.3%; Score 601; DB 1; Length 954;

Matches 155; Conservative 63; Mismatches 160; Indels 60; Gaps 10;

DB 10 QIKDDIOLFENCGGSGSVYKRWISQKKEVVKKL-LIKK-----EAIL 57

DB 92 EIPHELDLEETIVGGFGKVTALM--RGEVAVKARLPEDPAPTVAEQVCEARLF 149

Oy 58 SVLSHRNIIQFYGVILPEPNYGIYTEVASLGSLYDYNINSNSEMDMDHIM-TWATDVAK 116  
 Db 150 GALTGHPIIILARGLACTNPHCLTWEYARGALSRVLAGRRVP----PHVLVNAVQAR 205  
 Oy 117 GMHYLHMEAPVKVTHRDLSKSRNVI-----AADGVLIKIDFGASRFNHTTMSLVG 168  
 Db 206 GMNYLHNDAPVDIHRDLKSNIIILILEAIENHNLDVTLKIDFGLAREMHTTQMSAAG 265  
 Oy 169 TEPMAAPVIGSLPVSERCDPTYSYGVVLMEMLTREVPKSGLEGLOVAMLVYKNERLTIP 228  
 Db 266 TYAMAPPEVIRISLFSKSDVMSFGVLLMELTGVPRYREIDALVAAGVANNTLPLIP 325  
 Oy 229 SSCPSFAELLHQCEADAKRPFQKIIISLESMSNDTSLPDKCNSFLHNKAERCEIE 288  
 Db 326 STCPFPFARLLEECMDPPHGRPDGSLIKRLVEIGSALPQMLPSHSHIQEDMKLEIQ 385  
 Oy 289 ATLERKLTERDLSPKEOEL-----KERERLKMWEOKLTGOSNTPLLLPLAARMSSE 341  
 Db 386 HMFDDLFRYKEXELRREBELLRAGQEQFOEQLRREOELEREMDIIVERELHILMQL 445  
 Oy 342 SYFESKTESNSAEMSCQITATSNCEGCMPSLOAMLMNGCDLFSNNKAGAVHSGMQ 401  
 Db 446 SQEKPRVARKKGNFKRSRLKLREGSHISLPS-----GFER-----K 483  
 Oy 402 INMOAKONSSKTSKRG 419  
 Db 484 ITVQA-----SPTLDRKG 497

# RESULT 2

mixed-lineage protein kinase 1 - human  
 C:Species: Homo sapiens (man)  
 C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 17-Mar-1999  
 C:Accession: S32467; J00229  
 R:Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.  
 Eur. J. Biochem. 213, 701-710, 1993  
 A:Title: Identification of a new family of human epithelial protein kinases containing  
 A:Reference number: S32467; MUID:93238756; PMID:8477742  
 A:Accession: S32467  
 A:Molecule type: mRNA  
 A:Residues: 1-394 <D02>  
 C:Genetics:  
 A:Gene: GDB:MLK1  
 A:Cross-references: GDB:141921; OMIM:600136  
 A:Map position: 14q24.3-14q31  
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
 C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k  
 C:1-268/Domain: catalytic <CAT>  
 F:9-17/Region: protein kinase ATP-binding motif  
 F:289-310/Region: leucine zipper motif  
 F:324-345/Region: leucine zipper motif  
 F:354-368/Region: basic

Query Match 25.0%; Score 593; DB 2; Length 394;  
 Best Local Similarity 39.3%; Pred. No. 7, le-25;  
 Matches 133; Conservative 59; Mismatches 114; Indels 32; Gaps 5;

Oy 15 DLQFFENCQGSFGSVYRAKMSIDKEAVK-----KLLKLEKEAELISVL 62  
 Db 2 ELTLEELIIGIGFGKVRAPFAGID--EVAVKAARHPDEDISQITENYRQEKLPFAMLK 59  
 Oy 63 RNIIQFYGVILPEPNYGIYTEVASLGSLYDYNINSNSEMDMDHIMTWATDVAKMNY 122  
 Db 60 PNIIALRVCVCKEPLCLVMEFARGGPIARVLSGR--IPPDILVNAVOIARGMNY 116  
 Oy 123 MEAPVKVTHRDLSKSRNVI-----AADGVLIKIDFGASRFNHTTMSLVGFPWMA 174  
 Db 117 DEAVPIIHRDLKSNIIILIQKVENGDLSNKIKITDGLAREMHTTQMSAAGTYAMA 176  
 Oy 175 FEVIGSLPVSERCDPTYSYGVVLMEMLTREVPKSGLEGLOVAMLVYKNERLTIPSSCP 234

Db 177 DEVIRASWFSKSDVMSYGVLLMELTGEVPRGIDGLRVAAGVANNTLPISTCEP 236  
 Oy 235 FAELLHQCEADAKRPFQKIIISLESMSNDTSLPDKCNSFLHNKAERCEIATLERL 294  
 Db 237 FAKLMEDCQWNPDRSPFNIIIDQLTTIESGFEPMPKSFHCLQDNWKHEIOEMFDOL 296  
 Oy 295 KLERDLSPKEOEL-----KERERLKMWEOKLTGQ 325  
 Db 297 RAKEKELRWEELTRAAQKNOBELRRREGDLAR 334

# RESULT 3

mixed-lineage protein kinase (EC 2.7.1.-) 3 - human  
 N:Alternate names: protein kinase PTK1; protein kinase SPRK  
 C:Species: Homo sapiens (man)  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: A53800; I58395  
 R:Gallo, K.A.; Mark, M.R.; Scadden, D.T.; Wang, Z.; Gu, Q.; Godowski, P.J.  
 J. Biol. Chem. 269, 15092-15100, 1994  
 A:Title: Identification and characterization of SPRK, a novel src-homology 3 domain-cont  
 A:Reference number: A53800; MUID:94253068; PMID:8195146  
 A:Accession: A53800  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-847 <GAL>  
 A:Cross-references: GB:U07747; NID:9464027; PIDN:AAA19647.1; PID:9464028  
 R:Yang, Y.L.; Leung, I.W.; Heng, H.H.; Tsui, L.C.; Lassar, N.J.  
 Oncogene 9, 1745-1750, 1994  
 A:Title: MLK-3: identification of a widely-expressed protein kinase bearing an SH3 domai  
 A:Reference number: I58395; MUID:94239754; PMID:8183572  
 A:Accession: I58395  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-847 <RBS>  
 A:Cross-references: GB:J32976; NID:9488295; PIDN:AAA59859.1; PID:9488296  
 C:Genetics:  
 A:Gene: GDB:MLK3; PTK1; SPRK  
 A:Cross-references: GDB:134755; OMIM:600050  
 A:Map position: 11q13.1-11q13.3  
 C:Superfamily: mixed-lineage protein kinase 3; protein kinase homology; SH3 homology  
 C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein  
 F:115-383/Domain: SH3 homology <SH3>  
 F:143-131/Region: protein kinase homology <KIN>  
 F:403-424/Region: protein kinase ATP-binding motif  
 F:438-459/Region: leucine zipper motif  
 F:468-482/Region: basic

Query Match 24.5%; Score 581.5; DB 1; Length 847;  
 Best Local Similarity 37.4%; Pred. No. 6, 8e-24;  
 Matches 132; Conservative 60; Mismatches 122; Indels 39; Gaps 7;

Oy 13 FDDLQFFENCQGSFGSVYRAKMSIQ-----DKAVAVKLIKIEKEAELISVL 61  
 Db 114 FOELRELEVIGIGFGKVRAPFAGID--EVAVKAARHPDEDISV-TASVQEQEALFAMLA 172  
 Oy 62 HRNIIQFYGVILPEPNYGIYTEVASLGSLYDYNINSNSEMDMDHIM-TWATDVAKMNY 120  
 Db 173 HPNIIALKAACLEEPNCLVMEFARGGPIARVLSGR--IPPDILVNAVOIARGMNY 228  
 Oy 121 LHMEAPVKVTHRDLSKSRNVI-----IADGV-----LKIDFGASRFNHTTMSLVGFPW 172  
 Db 229 LHGALVPVTHRDLSKSNIIILQPIESDDMEHKLKIDTGLAREMHTTQMSAAGTYAM 288  
 Oy 173 MAPEVIGSLPVSERCDPTYSYGVVLMEMLTREVPKSGLEGLOVAMLVYKNERLTIPSSCP 232  
 Db 289 MAPEVIRASWFSKSDVMSYGVLLMELTGEVPRGIDGLRVAAGVANNTLPISTCEP 348  
 Oy 233 RSFAELLHQCEADAKRPFQKIIISLESMSNDTSLPDKCNSFLHNKAERCEIATLE 292  
 Db 349 EPEALQWADCAQDPHRRPDASTLQQLLEALQVLEMPRDPDSHMSQEKREIQGLFD 408





Db 246 YLHSHPT-VVHRDLKSNLVDKNWVVKVDFGLSRKHHHTYLSKSTAGTPEMNAEV 304  
 Qy 178 IQSLPSTCTDYSGVGLVEMLTREVPFKGLEIQAVALVKKERLITPSSCPSPFAE 237  
 Db 305 LRMEFANRCKCVYSGVGLVEMLTREVPFKGLEIQAVALVKKERLITPSSCPSPFAE 364  
 Qy 238 LHOQWEADAKRPSFKQIISLEMSNDTSLPDKNS 275  
 Db 365 IIRCEWQTEPHLRPSFTLQMSLKLQGLNLSNRPANT 401

## RESULT 7

JC5399  
 dual leucine zipper kinase (EC 2.7.-.-) - rat  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 07-Jul-1997 #sequence\_revision 18-Jul-1997 #text\_change 28-May-1999  
 C/Accession: JC5399  
 R/Matsumi, N.; Sarkar, G.; Shuto, T.; Maris, J.; Bronk, J.T.; Mizuno, K.; Bolander, M.E.; Biochem. Biophys. Res. Commun. 229, 571-576, 1996  
 Title: Identification of a dual leucine zipper kinase involved in rat fracture repair.  
 Reference number: JC5399; MUID:97127443; PMID:8954939  
 A/Accession: JC5399  
 A/Status: nucleic acid sequence not shown  
 A/Molecule type: mRNA  
 A/Residues: 1-888 <MAT>  
 C/Comment: This enzyme is involved in regulating cell function in the musculoskeletal system.  
 C/Genetics:  
 A/Genes: RDLK  
 C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
 C/Keywords: ATP; phosphotransferase  
 F/54-75, 89-98/Region: glycine-rich  
 F/156-454/Domain: kinase catalytic #status predicted <CAT>  
 F/156-404/Domain: protein kinase homology <KIN>  
 F/164-172/Region: protein kinase ATP-binding motif  
 F/421-449/Region: leucine zipper motif  
 F/472-500/Region: leucine zipper motif  
 F/557-888/Region: glycine-serine-proline rich #status predicted

Query Match 23.6%; Score 560.5; DB 2; Length 888;  
 Best Local Similarity 36.7%; Pred. No. 9.4e-23;  
 Matches 133; Conservative 73; Mismatches 131; Indels 25; Gaps 11;

Qy 10 QIKRDDLQPFENCGGSGSVYRAKVISODEKVAVKLLKTEKEAEI--LSVLSHNTIQ 67  
 Db 152 EAPFEIIDLQWVGSGAQAFLGRF--HGEVAVKVRDL--KETDIKHLKLNHPNIT 208  
 Qy 68 FYGVILBPPNYGIVTEYASLSGLYDINSNSEEMDMHMTWATDVAKGMYLHMEAPV 127  
 Db 209 FKGVCTQAPCYCLIMEFCAQOQVLEVAAGRPTVPSL--LVDMSGIGAGMYIHLH--- 263  
 Qy 128 KVIHRDLKSNVAVIADGVLCIDFGASR-FHNHTHSLVGTFFPMAPDEVIOQLPVSET 186  
 Db 264 KIIHRDLKSPMMLITVDVVVVISDFTSKSLSKSTKSPAGTAVAMAPAVIRNEPVSEK 323  
 Qy 187 CDTSYGVVLEMTREVPFKGLEIQAVALVKKERLITPSSCPSPFAELHOCWEAD 246  
 Db 324 VDIWSPGVVLEMTREVPFKGLEIQAVALVKKERLITPSSCPSPFAELHOCWEAD 383  
 Qy 247 AKRPSFKQIISLEMSNDT-SLPDKNSFLHNAKRECEIATELRLKLERDLSPKE 305  
 Db 384 PRNRPSPQIILHLDIASADVLTQF--TYFKSQAKREEVKLFKIKSEGCYLAHLE 441  
 Qy 306 QELKERERR-----LKMKEQKLTQESNTPLLLPLAARSEBSYFSK--TEESNSAEM 356  
 Db 442 EELVWRGRGFRHMLDIHEHYERKLERANN--LYMEINAMLQELTERELRRAQALER 499  
 Qy 357 SC 358  
 Db 500 RC 501

## RESULT 8

T16747  
 hypothetical protein R13f6.7 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 17-Mar-2000  
 C/Accession: T16747  
 R/Miller, N.  
 submitted to the EMBL Data Library, April 1994  
 A/Description: The sequence of C. elegans cosmid R13f6.  
 A/Reference number: Z18570  
 A/Accession: T16747  
 A/Status: preliminary; translated from GB/EMBL/DDBJ  
 A/Molecule type: DNA  
 A/Residues: 1-328 <MTL>  
 A/Cross-references: EMBL:U00046; NID:9470358; PID:9470364; PIDN:AAC47047.1; GSPDB:GN000.  
 A/Experimental source: strain Bristol N2; clone R13f6  
 C/Genetics:  
 A/Genes: CESP:R13f6.7  
 A/Map position: 3  
 A/Insertions: 20/3; 160/3; 222/2; 266/2  
 C/Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3

Query Match 23.0%; Score 545.5; DB 2; Length 328;  
 Best Local Similarity 41.2%; Pred. No. 2e-22;  
 Matches 112; Conservative 54; Mismatches 93; Indels 13; Gaps 5;

Qy 8 FVOIKPDDLOPFENCGGSGSVYRAKVISOD---KEVAVKLLKTEKEAEIISVLSHRN 64  
 Db 44 FPDIDRDIQVGHIGTGTAAPSGNWLTPDGSQRTALKVLEKEAEIISKIRHN 103  
 Qy 65 IIPGYVILBPPNYGIVTEYASLSGLYDINSNSEEM-----DMDHMTWATDVAK 116  
 Db 104 IIPGYVILBPPNYGIVTEYASLSGLYDINSNSEEM-----DMDHMTWATDVAK 163  
 Qy 117 GMYHMEAPVYKVIHRDLKSNVAVIADGVLCIDFGASR-FHNHT--THNSLVGTFFPMAP 175  
 Db 164 GIQVHDAVDVITIIHRDLKSNVAVIADGVLCIDFGASR-FHNHT--THNSLVGTFFPMAP 223  
 Qy 176 E-VIOSLPVSECTDYSGVGLVEMLTREVPFKGLEIQAVALVKKERLITPSSCPSP 234  
 Db 224 EMIIQSGLTATDVWVSGVGLVEMLTREVPFKGLEIQAVALVKKERLITPSSCPSP 283  
 Qy 235 FAEHLHOCWEADAKRPSFKQIISLEMSND 266  
 Db 284 LKQMSNCWMTPTDRANMRQIQEELNRLAGN 315

## RESULT 9

T48544  
 MAPK delta-1 protein kinase - Arabidopsis thaliana  
 N/Alternate names: protein F14F18.20  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
 C/Accession: T48544  
 R/Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft  
 submitted to the Protein Sequence Database, April 2000  
 A/Reference number: Z24490  
 A/Accession: T48544  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-886 <BEV>  
 A/Cross-references: EMBL:AL163812  
 A/Experimental source: cultivar Columbia; BAC clone F14F18  
 C/Genetics:  
 A/Map position: 5  
 A/Insertions: 148/3; 180/3; 327/3; 362/2; 568/3; 658/1; 671/3; 694/3; 728/2; 754/3; 810/3;  
 A/Note: F14F18.20

Query Match 22.3%; Score 529; DB 2; Length 886;  
 Best Local Similarity 43.3%; Pred. No. 4.5e-21;  
 Matches 107; Conservative 45; Mismatches 79; Indels 16; Gaps 5;

Qy 10 QIKRDDLQPFENCGGSGSVYRAKVISODEKVAVKLL-----KTEKEAEIISV 59  
 Db 107 QIKRDDLQPFENCGGSGSVYRAKVISODEKVAVKLL-----KTEKEAEIISV 59

Db 643 EIMMEDDQIGERIGISGYEYRLEM--NGTEVAVKKEFLDDFSGDALTOFKEIEIMLR 700  
Qy LSHRNIIQFVGVILEPPNYGIVTEYASIGSLYDINSNRSEMDMDHIMTAVAKGMH 119  
C/Accession: B87950  
R/anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo  
A/Reference number: A75000; MIMD:99069613; PMID:9951916  
A/Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C-el  
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an  
A/Accession: B87950  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-848 <STO>  
A/Cross-references: GB:chr\_1; PIDN:CAB06544.1; PID:g3876653; GSPDB:GN00019; CESP:F33E2.  
A/Map position: 1

Qy 238 LTHOCWE 244  
Db 879 IIRRCWQ 885

ULT 10  
763  
hypoetical protein F25P22.8 (imported) - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C/Accession: F96763  
R/Thellogis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Hultzer, L.  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MIMD:21016719; PMID:11130712  
A/Accession: F96763  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-1030 <STO>  
A/Cross-references: GB:AB051173; NID:g6692730; PIDN:AAF24036.1; GSPDB:GN00141  
C/Genetic: B  
A/Map position: 1

Query Match 21.5%; Score 510.5; DB 2; Length 1030;  
Best Local Similarity 38.0%; Pred. No. 5.2e-20;  
Matches 109; Conservative 54; Mismatches 93; Indels 31; Gaps 7;  
Qy 10 QIKRDDLOFPENCGSGFSGVYRAKMSODKEVAVKLL-----KIEKEAELISV 59  
Db 742 EIMMEDDQIGERIGISGYEYRLEM--NGTEVAVKKEFLDDFSGDALTOFKEIEIMLR 700  
Qy 60 LSHRNIIQFVGVILEPPNYGIVTEYASIGSLYDINSNRSEMDMDHIMTAVAKGMH 119  
Db 800 LSHRNIIQFVGVILEPPNYGIVTEYASIGSLYDINSNRSEMDMDHIMTAVAKGMH 119  
Qy 120 YLHMEAPVKVIRHDLKSNRVVIAADGVLCIDFGASRPHNHTTH---SLVGTFFMAPE 176  
Db 859 YLHMEAPVKVIRHDLKSNRVVIAADGVLCIDFGASRPHNHTTH---SLVGTFFMAPE 176  
Qy 177 VIQSLPVSETCDYTSYGVVLEWMLTREVFPKGLBGLQVAVLVEKNERLTISSCPRSPA 236  
Db 917 VLRNPAEKCDYTSYGVVLEWMLTREVFPKGLBGLQVAVLVEKNERLTISSCPRSPA 236  
Qy 237 ELLHOCWEADAKRPSFKQIISLESMS-----NDTSLP 270  
Db 977 DLISKWQTDKRLRPSFAIMASLRLQKPYTGSNIPRPVSSSLP 1023

RESULT 11  
B87950  
protein F33E2.2 (imported) - Caenorhabditis elegans

C/Species: Caenorhabditis elegans  
C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C/Accession: B87950  
R/anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo  
A/Reference number: A75000; MIMD:99069613; PMID:9951916  
A/Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C-el  
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an  
A/Accession: B87950  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-848 <STO>  
A/Cross-references: GB:chr\_1; PIDN:CAB06544.1; PID:g3876653; GSPDB:GN00019; CESP:F33E2.  
A/Map position: 1

Query Match 21.3%; Score 505; DB 2; Length 848;  
Best Local Similarity 30.4%; Pred. No. 8.3e-20;  
Matches 145; Conservative 69; Mismatches 159; Indels 104; Gaps 14;  
Qy 10 QIKRDDLOFPENCGSGFSGVYRAKMSODKEVAVKLLKIEKEAEI--LSVLSHRNIIQ 67  
Db 56 ELPDALSLEWLGSGGQGVFRQ--LENRIVAVKKNOL-KETIKHLRHLRQNIIE 112  
Qy 68 FYGVILEPPNYGIVTEYASIGSLYDINSNRSEMDMDHIMTAVAKGMHYLEAPV 127  
Db 113 FLYGVCSPPCYCIVMECSGQGLCTVAKSNRTITREL--FAQWKEIADGMHYLHQN--- 167  
Qy 128 KVIRHDLKSNRVVIAADGVLCIDFGASRPHN--HTTMSLVGTFFMAPEVIOSLPVE 185  
Db 168 KVIRHDLKSNRVVIAADGVLCIDFGASRPHN--HTTMSLVGTFFMAPEVIOSLPVE 185  
Qy 186 TCDYTSYGVVLEWMLTREVFPKGLBGLQVAVLVEKNERLTISSCPRSPAELHOCWEA 245  
Db 228 KVDYTSYGVVLEWMLTREVFPKGLBGLQVAVLVEKNERLTISSCPRSPAELHOCWEA 245  
Qy 246 DAKKRPSPFOI-----ISLESMSNDTSLPD-- 271  
Db 288 KGRNRPSPFHSIRHQLWEIPELPEMTEEWQAMDSYREFAKCIQYSTYTRDHGKSA 347  
Qy 272 -----KNSFLNKAEMRCEIEATLERLKLLEPDSFKEOELKERERLMMEOKL 322  
Db 348 FAMEEIORRHEQLNHIKDIRNMYEMKLRITNMVYDKLGCFTELKXESLEWMEKDL 407  
Qy 323 TEQ-----SNTP--LLPLAAMS--EESYFESTERS-----NSAEKSCQITATNSGEH 369  
Db 408 TEREQTHNQSPKAVAPRAQLRGYPNEGVDMSDDVQPCRGSPYRCSTNTSS----- 462  
Qy 370 GMPNLSQAMMLMGFDIFSMNKAQVAVHSGMQIMQAKONSSKTT--RRGKRVN 423  
Db 463 -----SGVQSSPFRQSSRSRAQQRIRSGAN 491

RESULT 12  
T20082  
hypoetical protein F33E2.2 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
C/Accession: T20082; T21703  
R/Lennard, N.  
submitted to the EMBL Data Library, April 1998  
A/Reference number: Z19221  
A/Accession: T20082  
A/Status: preliminary; translated from GB/EMBL/DDB3  
A/Molecule type: DNA  
A/Residues: 1-855 <WIL>  
A/Cross-references: EMBL:AL022593; PIDN:CAA18635.2; GSPDB:GN00019; CESP:F33E2.2  
R/Lennard, N.  
submitted to the EMBL Data Library, January 1997  
A/Reference number: Z19461

A:Accession: T21703  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-855 <W12>  
A:Cross-references: EMBL:Z84574; PIDN: CAB06544.2; GSPDB: GN00019; CESP: F33E2.2  
A:Experimental source: clone F33E2  
C:Genetics:  
A:Gene: CESP:F33E2.2  
A:Map position: 1  
A:Introns: 47/2; 213/2; 263/2; 366/3; 494/1; 547/3; 608/1; 826/3

Query Match            21.3%; Score 505; DB 2; Length 855;  
Best Local Similarity     30.4%; Pred. No. 8.3e-20;  
Matches    145; Conservative      69; Mismatches    159; Indels    104; Gaps    14;

OY    10 OIKPDLQPFENCGGSGPSVYRAKMI SODKEAYVKLIKIKEAEI--LSVLSHRNII Q 67  
     : |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
Db    56 ELIPFDISLEMLGGSSQGAVRGO--LENRTVAAYKVQL--KETELIKHLRLHRIIONIE 112  
  
OY    68 FYGVILLEPPNYGI VTEYASIGSLVDYINSNSEBMDHMATDPAVKMHYLMEAPV 127  
     I LFGVCSSPCYCIVMEYCSKGOLCTVLKSRNTITREL--FAQMVEKIADGMHYLHQN--- 167  
  
OY    128 KTIHNDLKARNVIADGYLTKICDRASRFNH--HTTHMSLVGTPTPWMAPEVIOGLPSFE 185  
     |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
Db    168 KVIHHDLPKSPLLISAEDEIKICDFETSMOKKMSTMMSPCGTYSWMAPENIKKQPCE 227  
  
OY    186 TCDDTSYGVALMEMLRVPFPGLEGLOVMLEVKENERTLTIPSCEPSFALLHCWEA 245  
     228 KDVVVSFGVVLVEMLTRERPANNIAOMATIIPVGNNILSLPMPEEARPKGLVLLIQCLSQ 287  
     246 DAKKRPSFOI-----ISTLESMSDTSLPD-- 271  
  
OY    288 KGRRNPFSFHIIQHWEIFKPBELEFMENTEEWOLAMPDSYREFACIOYPSPVTEDHGPRSA 347  
     272 -----KCNSPLHNKAEMRCETEATELERLKXLERDISFKEOELKERERLTKMWEOKL 322  
  
Db    348 FAMBEETIORKHQNLNIKDINMYEMKLKTRINKMTDKLGCFTELKIKESLAEMEKDL 407  
  
OY    323 TEQ----SNTP-LILLPIARWS-EESYFESKTEES---NSAEMSQQITATSNGEGH 369  
     408 TEREQWHNQNSKAAVAAPAOLRGYPNEGYYDDMSDEDQVRGRSPYRCSNITS----- 462  
  
OY    370 GGNPBLQAMMLMGPDIFSNMKAGAMHSQMGINNAKNSKTTS---KRGKKRYN 423  
     463 ----\*-|||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
     -SGVQSSPFRQSRSRSSAQGOTRRSEGAN 491

MULT 13  
576  
probable protein kinase TCTR2 - tomato  
C:Species: Lycopersicon esculentum (tomato)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 08-Oct-1999  
C:Accession: T06576  
R:Hackett, R.M.  
submitted to the EMBL Data Library, March 1998  
A:Reference number: Z15770  
A:Accession: T06576  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-982 <HA>  
A:Cross-references: EMBL:AJ005077; NID:e1296722; PIDN:CAA0634.1; PID:e1296723  
A:Experimental source: cultivar Alisa crataegus  
C:Genetics:  
A:Gene: TCTR2

Query Match            20.8%; Score 495; DB 2; Length 982;  
Best Local Similarity     38.9%; Pred. No. 3.3e-19;  
Matches    110; Conservative      49; Mismatches    104; Indels    20; Gaps    7;

OY    10 OIKPDLQPFENCGGSGPSVYRAKMI SODKEAYVKLIKIKEAEI-----KIKEAEILSV 59  
     : |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
Db    695 ELTPVEDLVIGEIRIGGSIGEVYHADW-NGETVAKKFLDQDFGAALAFPERVRIOR 752

```

QY 60 ILSHRIIOGYVITLPEPNVGIWTEVASIGSLYDYINSNRSEEMDMHITWMTDVGKMH 119
Db 753 LHPHNVVRMGWILTPPHLSITTEFLPRGSLYRIIHRHPLF QIDEROKIKMLDVAKGD 811
QY 120 VYHMEAPVKVIRHDLKSRNVVIADGVLTICDYGASRPHNHT -THMSLVGTFPMWAP 177
Db 812 CLHTSNPT -IVHRDLKSPMLVLDTDMMNVKCDYGLSRLKHNFTLSKSTAGPEWMAPEV 870
QY 178 IOSLVSSECDYVSGVVMWMLTREVPKGLGEGVAMLVYKKEKRLIIPSSCPRSF 237
Db 871 LNPSNSNEKCDIYSGVVIMELATLRLPWSGNPMQVAGVAPQNRLEIPELDEIVAR 930
QY 238 LHHQCEADAKRPSFQKQIISILSMSNDTSLP---DKCNSFL 277
Db 931 IWECCQMDTPNLRPSFAQLTVALTPLOK -LVIPAVVDQNSRL 972

RESULT 14
T8400
serine/threonine-protein kinase ctrl - Arabidopsis thaliana
N.Alternate names: protein P17C15.150
C.Species: Arabidopsis thaliana (mouse-ear cress)
C.Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 17-Nov-2000
C.Accession: I48400; A45178
R.Bevan, M., Pohl, T., Weizenegeger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.
submitted to the Protein Sequence Database, March 2000
A.Reference number: I24492
A.Accession: T8400
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-821 <BEV>
A.Cross-references: EMBL:AL162506
A.Experimental source: cultivar Columbia; BAC clone P17C15
R.Kieber, J.J.; Rothenberg, M.; Roman, G.; Feldmann, K.A.; Ecker, J.R.
Cell 72, 427-441, 1993
A.Title: CTR1, a negative regulator of the ethylene response pathway in Arabidopsis, en
A.Reference number: A45178; MUID:93161417; PMID:8431946
A.Contents: Columbia
A.Accession: A45178
A.Status: preliminary; not compared with conceptual translation
A.Molecule type: DNA
A.Residues: 1-468, 470-821 <KIE>
A.Note: sequence extracted from NCBI backbone (NCBI:124878)
C.Genetics:
A.Map position: 5
A.Introns: 216/3; 317/2; 364/3; 399/2; 468/1; 560/1; 573/3; 596/3; 630/2; 658/3; 691/3;
A.Note: F17C15.150
C.Keywords: ATP
P:548-812/Domain: protein kinase homology <KIN>
P:551-565/Region: protein kinase ATP-binding motif

Query Match 20.5%; Score 488; DB 2; Length 821;
Best Local Similarity 40.7%; Pred. No. 6.5e-19;
Matches 109; Conservative 44; Mismatches 39; Indels 16; Gaps 5;

QY 9 VOIKRDDIIOFENCGGSGFSGYRAKWTISDPEKVAVKKLKE-----KEAEILS 58
Db 544 MDIPWCNDINIEKIGAGSFGTVHRAEMHGSD--VAVKILMEDDFAHARVNEFLREVALIK 601
QY 59 VLSHNNIIOFGVILPEPNVGIWTEVASIGSLYDYIN -SNSSEEMDMHIMWATDVAKG 117
Db 602 RLHSHNVITVPMGAVYQPNLSLIVTEYLSRGSILRLHHSGARQEDERRRLSMADVAKG 661
QY 118 MHTLMEAPVKVIRHDLKSRNVVIADGVLTICDYGASRPHNHT -THMSLVGTFPMWAP 175
Db 662 MNYLTNRP -PIVHRDLKSPMLVLDTKYVAVKCDYGLSRLKASTFLSSKSAAGTEWMAPE 720
QY 176 EYIOSLVSSECDYVSGVVMWMLTREVPKGLGEGVAMLVYKKEKRLIIPSSCPRSF 235
Db 721 EYLRPEBSNESDYSFGVILMELATLLOQKONLPAQVAAVAGKCKRLEIPELRLNPQV 780

```

```

RESULT 14
T88400 serine/threonine-protein kinase ctrl - Arabidopsis thaliana
N/Alternate names: protein F17C15.150
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 17-Nov-2000
C/accession: T88400; A45178
R/Beran, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.
submitted to the Protein Sequence Database, March 2000
A/reference number: Z24492
A/accession: T8400
A/status: preliminary
A/molecule type: DNA
A/residues: 1-821 <BBV>
A/cross-references: EMBL:AL162506
A/experimental source: cultivar Columbia; BAC clone F17C15
R/Kieber, J.J.; Rothenberg, M.; Roman, G.; Feldmann, K.A.; Ecker, J.R.
Cell 72, 427-441, 1993
A/title: CTRL, a negative regulator of the ethylene response pathway in Arabidopsis, en
A/reference number: A45178; MUID:93161417; PMID:8431946
A/content: Columbia
A/accession: A45178
A/status: preliminary; not compared with conceptual translation
A/molecule type: DNA
A/residues: 1-468,470-821 <KIE>
A/name: sequence extracted from NCBI backbone (NCBI:124878)
A/map position: 5
C/genetics:
A/introns: 216/3; 317/2; 364/3; 399/2; 468/1; 560/1; 573/3; 596/3; 630/2; 658/3; 691/3;
A/name: F17C15.150
C/superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol
C/key words: ATP
F/549-812/Domain: protein kinase ATP-binding motif
P/557-565/Region: protein kinase ATP-binding motif

Query Match      20.5%   Score 488; DB 2; Length 821;
Best Local Similarity 40.7%; Pred. No. 6.5e-19;
Matches 109; Conservative 44; Mismatches 99; Indels 16; Gaps 5;

QY 9 VOIKEDDLOFFENCGSGFSYYRAKWISODKEVAVKLLKE-----KEAIIIS 58
DB 544 MDIPMCDINTIHEKIGAGSFGTYHRAEMHGSD--VAVKILMEDPFHAERVNETLRVALAIK 601
QY 59 VLSHRNIIOFYGVILLEPPNYGIVTEYASLGSLVDYIN-SNRSEENDMDHMIMATDVAKG 117
DB 602 RLRHNNIVLFMGAVIQPNPLSIIVEYLIRGSLRYLLRHLSGARSEQDERRRRLSMADVAKG 661
QY 118 MYHLMEAPVKYIRHDLSKNRVVIADGVLKICDPGARFRFHNT--THMSLVGTFPWAP 175
DB 662 MYVLNNRP-PIVHRDLKSPNLIDKKYTVKVCDFGLSRKLASTLSKSAAGTPEWMAP 720
QY 176 EVIOSLPSETCDTTSYGVMLEMLTREVPFKGLEGLQVAMLVKENRELTTPSCPRSF 235
DB 721 EYLRBPSNEKSDVYSFGVILLIELTLTQQPMONLNPQOVVAAGVKCKRKLPIRNINPOV 780

```

QY 236-261LHOCWEADAKKRPSPKQIISILESM 263  
 Db 781 AAIIEGCTWNEPWRKRPSPFATIMDLRPL 808

## RESULT 15

JCS955  
 transforming growth factor-beta activated kinase (EC 2.7.-.-) 1a - human  
 C/Species: Homo sapiens (man)  
 C/Date: 16-Jul-1999 #sequence\_reviseion 16-Jul-1999 #text\_change 21-Jul-2000  
 C/Accession: JCS955  
 R/Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.  
 Biochem. Biophys. Res. Commun. 243, 545-549, 1998  
 A/Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind  
 A/Reference number: JCS955; MUID:98153801; PMID:9480845  
 A/Accession: JCS955  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-579 <SAK>  
 A/Cross-references: DDBJ:AB009356; NID:92924623; PIDN:BAA25025.1; PID:92924624  
 A/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
 C/Keywords: phosphotransferase

Query Match 20.3%; Score 483; DB 2; Length 579;  
 Best Local Similarity 30.2%; Pred. No. 8.1e-19;  
 Matches 131; Conservative 82; Mismatches 151; Indels 70; Gaps 18;

QY 7 SFVQIKFPDDLOFFENCSSFGSVYRAKVISODKEVAVKLLKTEKEAE-----IL 57  
 Db 27 NFEIDYKEIVEEVNGAIFGVCKAKM--RAKDVAIK--QIESESEKAFVELRQL 81  
 QY 58 SVLSHRNIIQFYGVILEPPNYGIYEVASLSGLYDINSRSEEM--DMDHMTWATDV 114  
 Db 82 SRVNHPNIVKLYGACLN--VCLWVEVAGGSLVNVLHG--AEPLPYTAHAAMSWCLQC 137  
 QY 115 AKGHVYLMEMPAVYIHRDLKSRNVVIADG-VLKICDFG-ASRPHNTTMSLVGTFPM 172  
 Db 138 SQGVAVYLSMQPKALIHRLKPNLLVAGGTVLKICDFGTACDIQTMTNNK--GSAAM 195  
 QY 173 MAPENVIGSLPSETCDTYSYGVLMEMLTREVPFKGLEG--LOVAMLVVEKNERLTIPSS 230  
 Db 196 MAPEVPESGNSSEKCDVPSWGIILMEVITRKPFDEIGGPAFRIMW-AVANGTRPPLIKN 254  
 QY 231 CPRSFABILLHOCWEADAKKRPSPKQIISIL-----ESMSNDT- 267  
 Db 255 LFKPIESIMTRCWSKDPQRPSPMEIVKIMTHLMRYPPGADEPLQYPCQYSDGQNSGAT 314  
 QY 268 ---SLPDKNSFLNKAERCE-IEATLERLKLRLDSFKOEIKERERLKMMEQKLT 323  
 Db 315 STGSFMDIASTNTSNKSDPTNMEQVPATMDTKRLSKLLKNQAKQSSGRLSLGASRGS 374  
 QY 324 EQSNTPLLLPLAARMSESYFESKTESNSAEMSCQITATSNGEHGNNPSLOAMLMGF 383  
 Db 375 SVESLP---PTS-----EGKMSADMSEIEIRIATTT-GNGQPRRSIIDLTVTGT 421  
 QY 384 --GDIFSNKAGAV 395  
 Db 422 EPGQVSSRSRSPSV 435

Search completed: May 1, 2003, 20:38:37  
 Job time : 37.2873 secs



GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:13:36 ; Search time 22.8452 Seconds  
(without alignments)  
826.070 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 2375  
Sequence: 1 MSSLGASFWQIKRFDLQFFE.....GDDDDDDGSEEDNDNDNSE 455

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	601	25.3	954	1 M3KA_HUMAN	Q02779 homo sapien
2	593	25.0	394	1 M3K9_HUMAN	P80192 homo sapien
3	565.5	23.8	888	1 M3KC_HUMAN	Q12852 homo sapien
4	565.5	23.7	888	1 M3KC_MOUSE	Q60700 mus musculu
5	563.5	23.7	888	1 M3KC_RAT	Q63796 rattus norv
6	488	20.5	821	1 CTR1_ARATH	Q05609 arabidopsis
7	483	20.3	579	1 M3K7_MOUSE	Q62073 mus musculu
8	471.5	19.9	606	1 M3K7_HUMAN	Q43318 homo sapien
9	434	18.3	746	1 ABL_MLVAB	P00521 abelson mur
10	410	17.3	1130	1 ABL1_HUMAN	P00519 homo sapien
11	405	17.1	1123	1 ABL1_MOUSE	P00520 mus musculu
12	405	16.9	1520	1 ABL1_DROME	P02684 homo sapien
13	402	16.9	1520	1 ABL1_DROME	P00522 drosophila
14	401.5	16.2	439	1 ABL_FSVAY	P18160 dictyostell
15	385.5	16.2	536	1 FYN_XIPHE	P10447 feline sarc
16	385.5	16.2	536	1 FYN_XIPHE	P27446 xiphophorus
17	379.5	15.9	505	1 FRK_HUMAN	P24665 homo sapien
18	378.5	15.9	536	1 FRK_DROME	P33104 drosophila
19	378.5	15.9	536	1 FYN_XENLA	P13406 xenopus lae
20	376.5	15.7	536	1 FYN_HUMAN	P06241 homo sapien
21	374	15.7	410	1 KYK2_DICDI	P18161 dictyostell
22	369.5	15.6	1196	1 ABL1_CAEEL	P03949 caenorhabdi
23	365.5	15.4	535	1 YRK_CHICK	Q02977 gallus gall
24	359.5	15.1	587	1 SRC_MOUSE	P19688 mus musculu
25	358.5	15.1	675	1 SRC_AVIS2	P15054 avian sarco
26	358	15.1	822	1 BMK_HUMAN	P18183 homo sapien
27	358	15.1	822	1 FGRI_HUMAN	P11362 homo sapien
28	358	15.1	822	1 FGRI_HUMAN	P11362 homo sapien
29	357	15.0	552	1 FGRI_DROME	P00528 drosophila
30	357	15.0	819	1 FGRI_CHICK	P18044 gallus gall
31	355.5	15.0	517	1 FGR_MOUSE	P14224 mus musculu
32	355.5	15.0	531	1 SRC1_XENLA	P13115 xenopus lae
33	355.5	15.0	531	1 SRC2_XENLA	P13116 xenopus lae

34	354	14.9	497	1 SPK1_DUGTI	P42687 dugesia cig
35	353.5	14.9	532	1 SRC_CHICK	P00523 gallus gall
36	353	14.9	822	1 FGRI_RAT	P00523 rattus norv
37	352	14.8	806	1 CERK_CHICK	P18460 gallus gall
38	351.5	14.8	557	1 SRC_AVIS1	P14085 avian sarco
39	351.5	14.8	757	1 HT16_HYDAT	P53356 hydra atten
40	350.5	14.8	527	1 TXK_HUMAN	P42681 homo sapien
41	350.5	14.8	533	1 FYN_CHICK	P00587 gallus gall
42	350	14.7	806	1 FGRI_HUMAN	P22607 homo sapien
43	350	14.7	1426	1 EGFR_DROME	P04412 drosophila
44	349.5	14.7	535	1 SRC_HUMAN	P12931 homo sapien
45	349.5	14.7	568	1 SRC_AVIS3	P14084 avian sarco

## ALIGNMENTS

RESULT 1  
M3KA\_HUMAN STANDARD; PRT; 954 AA.  
ID M3KA\_HUMAN Q02779; Q12761; Q14871;  
AC Q02779; Q12761; Q14871;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Mitogen-activated protein kinase kinase kinase 10 (EC 2.7.1.37)  
DE (Mixed lineage kinase 2) (Protein kinase MST)  
GN MAP3K10 OR MLK2 OR MST.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=96128179; PubMed=8536694;  
RA Dorow D.S., Devereux L., Tu G.F., Price G., Nicholl J.K.,  
RA Sutherland G.R., Simpson R.V.;  
RT "Complete nucleotide sequence, expression, and chromosomal  
RT localisation of human mixed-lineage kinase 2.";  
RL Eur. J. Biochem. 234:492-500(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=95249256; PubMed=7731697;  
RA Katoh M., Hirai M., Sugimura T., Terada M.;  
RT "Cloning and characterization of MST, a novel (putative)  
RT serine/threonine kinase with SH3 domain.";  
RL Oncogene 10:1447-1451(1995).  
RN [3]  
RP SEQUENCE OF 244-480 FROM N.A.  
RC TISSUE=Colon epithelium;  
RX MEDLINE=93238756; PubMed=8477742;  
RA Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;  
RT "Identification of a new family of human epithelial protein kinases  
RT containing two leucine/isoleucine-zipper domains.";  
RL Eur. J. Biochem. 213:701-710(1993).  
CC -/- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -/- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND SKELETAL MUSCLE.  
CC -/- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC MAP KINASE KINASE SUBFAMILY.  
CC -/- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
DR EMBL: X90846; CAA62351.1; -;  
DR EMBL: Z48615; CAA88531.1; -;  
DR PIR: S32468; S32468.

HSP, pL1362; IFGK.  
DR Genew; HGNC:6849; MAP3K10.  
DR MIM; 600137;  
DR InterPro; IPR000719; Euk\_Pkinase.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR004040; Sry\_Pkinase.  
DR InterPro; IPR002290; Ser\_Thr\_Pkinase.  
DR InterPro; IPR001245; Tyr\_Pkinase.  
DR Pfam; PF00018; SH3; 1.  
DR Pfam; PF00069; Pkinase; 1.  
DR PRINTS; PR00452; SH3DOMAIN.  
DR PRINTS; PRO0109; TYRKINASE.  
DR ProDom; PD000001; Euk\_Pkinase; 1.  
DR ProDom; PD000066; SH3; 1.  
DR SMART; SM00326; SH3; 1.  
DR SMART; SM00221; STYKC; 1.  
DR PROSITE; PS01017; PROTEIN KINASE\_ATP; 1.  
DR PROSITE; PS01018; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PSS0011; PROTEIN\_KINASE\_DOM; 1.  
DR Transferrase; Serine/threonine-protein kinase; Tyrosine-protein kinase;  
ATP-binding; SH3 domain.  
FT DOMAIN 2  
FT DOMAIN 16 81 POLY-GLU.  
FT NP\_BIND 98 360 SH3.  
FT BINDING 104 112 ATP (BY SIMILARITY).  
FT ACT\_SITE 125 125 ATP (BY SIMILARITY).  
FT DOMAIN 222 222 BY SIMILARITY.  
FT DOMAIN 384 405 LEUCINE-ZIPPER 1 (BY SIMILARITY).  
FT DOMAIN 419 440 LEUCINE-ZIPPER 2 (BY SIMILARITY).  
FT DOMAIN 449 463 ARG/LYS-RICH (BASIC).  
FT CONFLICT 462 464 SRL -> AV (IN REF. 2).  
FT CONFLICT 465 480 LKEGGSHISLPSEGF -> AQAAGRQHPALML (IN REF. 3).

G->S (IN REF. 2).  
G->R (IN REF. 2).  
V->A (IN REF. 2).

Query March 25.3%; Score 601; DB 1; Length 954;  
Best local similarity 35.4%; Pred. No. 7, 5e-32;  
Matches 155; Conservative 63; Mismatches 160; Indels 60; Gaps

10 QIKEDDLQFFENCGSGSFVSYPRAVKIISODKEAVAKL-LKIEK-----EAEIL 57  
::|::|||::|::|::|::|::|:  
Db 92 EIPEHLEQLHEIIIGVGCKYVRALM--RGESVAVKARLDDEKDPAVTAEQCCEARLF 149  
58 SVLSHRNIIOFYGVILEEPNYGIIVETASGLSYLYINSNREEMDMOHIM-TWATDIAK 116  
150 GALCPHNITIALAGACLNPHCLTVEIRVARGSLRYLAGRVVP---PVLVNMAYVAR 205  
Db 117 GMHYLMNEAPYKVVIHRDKSRNVVI-----AADVLKICDFGASRFNNHTMSIWG 168  
206 GNMYINDNPAPVIPIHRLDKSINITILLERAIENHLADTVLKIDTFGLAREMKHTTKMSAA 265  
Qy 169 TTPMAAPEVISQLPSFHCPTITSYGVMLEMTRIENVPPREGLEGLOVAMLYVKNERLTIP 228  
266 TVAMAFAEVTRISTLFSKSDVWSFGVLLMWELLTGFEVPREIDALAIVAANNTKLTLRP 325  
Qy 229 SSCPRSFAELHLQCCMEADAKKRRPFKOJLIISLESNSNDTSLPDKCNSTFLHNKAERCETE 288  
Db 326 STCPEFPFARLLECCHDPBPGRPDGSILKRLEVIESALTFCMPLEPSHSLOEDWKRIEQ 385  
Qy 289 ATLIERLIKLERDSLSCREQEL-----KERERRLKMEEOKLTEQSNTPLLPFLAARMSEE 341  
Db 386 HMFDDLRTYEKEIELAREEBELIRAABOBFQEOLCTRRREEELAREMDIVERELHLMCOL 445  
Qy 342 SYVESKTSESNAEMSQCOTATSNGEGHMNSTLAMMLMGRCGDI FSNMKAGA VAWHSGNQ 401  
Db 446 SOEKFRVARRKKNFKRSRLIKLRBGESHSLPS-----GEFH-----K 483  
Qy 402 INMOAKONSKITTSKRQG 419  
::|:

Db	494	ITVOA----	SPTLDRKG	497
RESULT	2			
M3K9	HUMAN	STANDARD;	PRT;	394 AA.
ID	M3K9_HUMAN			
AC	P80192;			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Mitogen-activated protein kinase kinase kinase 9 (EC 2.7.1.-) (Mixed lineage kinase 1) (Fragment).			
DE	MAP3K9 OR MLK1 OR PRKEL.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
NC	NCBI_TaxID=9606;			
LN	[1]			
RP	SEQUENCE FROM N.A.			
RP	TISSUE=Colon epithelium;			
RC	MEDLINE=93238756; PubMed=8477742;			
RA	Dorow D.S., Devereux L., Dietzsch E., de Krester T.;			
RT	"Identification of a new family of human epithelial protein kinases containing two leucine/isoleucine-zipper domains.";			
RT	Eur. J. Biochem. 213:701-710(1993).			
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN EPITHELIAL TUMOR CELL LINES OF COLONIC, BREAST AND OESOPHAGEAL ORIGIN.			
CC	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.			
CC	MAP KINASE KINASE KINASE SUBFAMILY.			
CC	PIR; S32467; S32467.			
DR	PIR; J00229; J00229.			
DR	HSSP; P12931; IFMK.			
DR	Genew; HGNC:6861; MAP3K9.			
DR	MIM; 600136; -.			
DR	InterPro; IPR000719; Euk_Pkinase.			
DR	InterPro; IPR002290; Ser_Thr_Pkinase.			
DR	InterPro; IPR001245; Tyr_Pkinase.			
DR	Pfam; PF00069; Pkinase; 1.			
DR	ProDom; PD000001; Euk_Pkinase; 1.			
DR	SMART; SMO0219; TyrcK; 1.			
DR	PROSITE; PS00107; PROTEIN KINASE ATP; 1.			
DR	PROSITE; PS00108; PROTEIN KINASE_ST; 1.			
DR	PROSITE; PS50011; PROTEIN KINASE_DOM; 1.			
KX	Transferrase; Serine/threonine-protein kinase; Tyrosine-protein kinase; ATP-binding.			
KW	ATP-binding.			
FT	NON TER	1		
FT	DOMAIN	3	271	
FT	NP_BIND	9	17	
FT	BINDING	30	30	
FT	ACT SITE	127	127	
FT	DOMAIN	289	310	
FT	DOMAIN	324	345	
FT	DOMAIN	354	368	
FT	SEQUENCE	394 AA;	44975 MW;	DBE40B7D31047P08 CRC64;
QY	Query Match	25.0%;	Score 593;	DB 1; Length 394;
QY	Best Local Similarity	39.3%;	Pred. No. 8.2e-32;	
QY	Matches 133;	Conservative 59;	Mismatches 114;	Indels 32; Gaps 5;
Db	15	DIQFENGCGGSGFSGVYAKWISDQKEVAVK-----	KLKTKEKAEILSLVLSH	62
Db	2	ELTLEELIIGIGFGFVYAFWIGD--EVAVNAARHDPDEDISQTIENVRQEKIFAMLIK	H	59
QY	63	RNIIOFVAVILLEPPYVGLVTEFASLGSLYDYINSRSEBMDHIMTAVDAKGMATLH		122
Db	60	EVITIALRGVCLKEPMLCLWMEFARGGPIILRVLSGR--	IPPDILVMAVQILARGMATLH	116
QY	123	MEAPKVIHRDLKSNVY-----AADVLKICDPGASRFNNHTTMSLVGTFPWA		174
Db	117	DEALVPIIHRDLKSNIIILQKVENGDLSNKKIKTKTDGLAREMHRTTKMSAAGTYAWA		176
QY	175	PEVIGSLPVSCTDYISGVGVIMEMLTREVPFKGLGSLQVAMLVKEKRELLTPSCRS		234



Db 177 PEVIRASMFSGSDVSWISGVLWELLTGEVFRGIDGLRVAVGVANAKLALPISTCEP 236  
 QY 235 FAELHOCWEADAKKRSFKQIISLESMSNDTSLPDKCNSFLNKAEWCEIATEIRL 234  
 Db 237 FAKMECDWMDPHSRPSFTNILDQLTTIESGPFEMPDSFHCLODNNWHEIQEMFDOL 296  
 QY 295 KKLERDLSFEQEL-----KERERLKKMEOKLREQ 325  
 Db 297 RAKEKELTWEELTRAALOOKOEELLRRRDEOLAER 334

RESULT 3  
 M3KC\_HUMAN STANDARD; PRT; 859 AA.  
 AC Q12852;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Mitogen-activated protein kinase kinase kinase 12 (EC 2.7.1.37)  
 DE (leucine-zipper protein kinase) (ZPK).  
 GN MAP3K12 OR ZPK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Tetradocarcinoma;  
 RX MEDLINE=94311945; PubMed=8037767;  
 RA Reddy U.R., Pleasure D.;  
 RT "Cloning of a novel putative protein kinase having a leucine zipper  
 RT domain from human brain";  
 RT Biochem. Biophys. Res. Commun. 202:613-620(1994).  
 RT - FUNCTION: May be an activator of the JNK/SAPK pathway.  
 RT Phosphorylates beta-casein, histone 1 and myelin basic protein in  
 RT vitro.  
 CC -1 CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -1 COFACTOR: Magnesium.  
 CC -1 SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By  
 CC similarity).  
 CC -1 TISSUE SPECIFICITY: Highly expressed in brain and kidney.  
 CC -1 PIM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol  
 CC under basal conditions and dephosphorylated when membrane-  
 CC associated (By similarity).  
 CC -1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MAP KINASE KINASE KINASE SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U07358; AAA67343.1; -  
 CC HSPB; P12931; IFMK.  
 CC GeneW; HGNC:6851; MAP3K12.  
 CC MIM; 600447; -  
 CC InterPro; IPR000719; Euk\_Pkinase.  
 CC InterPro; IPR004040; STY\_Pkinase.  
 CC InterPro; IPR002290; Ser\_thr\_Pkinase.  
 CC Pfam; PF00069; Pkinase; 1.  
 CC ProDom; PD000001; Euk\_Pkinase; 1.  
 CC SMART; SM00221; STYK; 1.  
 CC DR PROSITE; PS00107; PROTEIN KINASE ST; 1.  
 CC DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 CC TRaNsEtase; Serine/threonine-protein kinase; ATP-binding;  
 CC Phosphorylation; Magnesium; Membrane.  
 CC FT DOMAIN 125 366  
 CC NP BIND 131 139  
 CC BINDING 152 152  
 CC ATP (BY SIMILARITY).  
 CC ATP (BY SIMILARITY).

FT ACT SITE 236 236 BY SIMILARITY.  
 FT DOMAIN 665 668 POLY-PRO.  
 FT DOMAIN 720 725 POLY-GLU.  
 SQ SEQUENCE 859 AA; 93188 MW; 0B520972C5C6F05 CRC64;

Query Match 23.8%; Score 565.5; DB 1; Length 859;  
 Best Local Similarity 37.8%; Pred. No. 1,4e-29;  
 Matches 135; Conservative 71; Mismatches 116; Indels 35; Gaps 13;

QY 10 QIKEDDLOFPENCGSGSGSVYRAKWIQDEVAFKKLTKIEKAEI--LSVSHRITIQ 67  
 Db 119 EVPFEEILDQWGSAGANVFLGRF--HGEVAVKVRDL-KETDIKHLRKLGNPIIT 175  
 QY 68 FYGVILEPPNATGYTEVASLGSQLYDYINSRSEMDMDHMTATDVAKGHHYLMKAPV 127  
 Db 176 FKVGCTAPCCCLIMERCAGQQLYEYVLRAPRPVPSL--LVDSMSGIAGANTYLHLH--- 230  
 QY 128 KVIRDLKSRNVVIAAGVAKICDPGASR--FHHHTHMSLVGTFPMMAPIVQSLSVSET 186  
 Db 231 KIHRDLKSPMLITVDYVKISDFGTSKELSDKSTMSFAGTVAMWAPVINEPVSSEK 230  
 QY 187 CDITYSYGVILWMLTREVPFKGLEQLQVAMLVYEKNERLTPSSCPSPFAELHOCWEAD 246  
 Db 291 VDWSFGVIMELLTGERIPYKQDVSSALIWGVSNSLHLPSPSCDYGFKILRQCNK 350  
 QY 247 AKKRSPFKQIISLESMSNDT-SLPDKCNSFLNKAEWCEIATEIRL-----KLE 298  
 Db 351 PRNPSFRQILHLHDASAVISTPOE--TYFSGAEMVEVLAHEKIKSEGTCLHRE 408  
 QY 299 RDLSEFQEQE-----LKERERLKKMEOKLREQSNTPLLLPIARMESESYFSEKTE 350  
 Db 409 EELVWRREELRHALDIRH---YERKLERANN--LYNELNALMLQ---LEIKERE 456

RESULT 4  
 M3KC\_MOUSE STANDARD; PRT; 888 AA.  
 ID M3KC\_MOUSE  
 AC Q60700; P70286;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Mitogen-activated protein kinase kinase kinase 12 (EC 2.7.1.37)  
 DE (leucine-zipper protein kinase) (ZPK) (Dual leucine zipper bearing  
 DE kinase) (DLK).  
 GN MAP3K12 OR ZPK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CD-1; TISSUE=Brain;  
 RX MEDLINE=95074107; PubMed=7983011;  
 RA Holzman L.B., Merritt S.E., Fan G.;  
 RT "Identification, molecular cloning, and characterization of dual  
 RT leucine zipper bearing kinase. A novel serine/threonine protein kinase  
 RT that defines a second subfamily of mixed lineage kinases";  
 RT J. Biol. Chem. 269:30808-30817(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ICR X Swiss Webster; TISSUE=Brain;  
 RX MEDLINE=96365388; PubMed=8769565;  
 RA Blouin R., Beaudoin J., Bergeron P., Nadeau A., Grondin G.;  
 RT "Cell-specific expression of the ZPK gene in adult mouse tissues";  
 RT DNA Cell Biol. 15:631-642(1996).  
 RN [3]  
 RP PHOSPHORYLATION AND MUTAGENESIS OF LYS-185 AND GLU-192.  
 RX MEDLINE=96279269; PubMed=8663324;  
 RA Mata M., Merritt S.E., Fan G., Yu G.G., Holzman L.B.;  
 RT "Characterization of dual leucine zipper-bearing kinase, a mixed  
 RT lineage kinase present in synaptic terminals whose phosphorylation  
 RT state is regulated by membrane depolarization via calcineurin";  
 RT J. Biol. Chem. 271:16688-16696(1996).

CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.  
 CC Phosphorylates beta-casein, histone 1 and myelin basic protein in  
 CC vitro.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -1- COFACTOR: Magnesium.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.  
 CC -1- TISSUE SPECIFICITY: Expressed in brain, kidney, lung, heart,  
 CC testis, gastrointestinal tract, stomach, liver and pancreas.  
 CC Within the nervous system, predominantly expressed in neurons and  
 CC enriched in synaptic terminals.  
 CC -1- PFM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol  
 CC under basal conditions and dephosphorylated when membrane-  
 CC associated.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MAP KINASE KINASE SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; U14636; AAA57280.1; -  
 CC DR EMBL; U23789; AAB17123.1; -  
 CC DR HSSP; P12931; IFMK.  
 CC MGP; MG1.1346881; Map3K12.  
 CC DR InterPro; IPR000719; Euk\_pkinase.  
 CC DR InterPro; IPR004040; STY\_pkinase.  
 CC DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 CC DR Pfam; PF00069; Pkinase; 1.  
 CC DR ProDom; PD000001; Euk\_pkinase; 1.  
 CC DR SMART; SM00221; STYKC; 1.  
 CC DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE\_NEG.  
 CC DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 CC DR PROSITE; PS50011; Serine/threonine-protein kinase; ATP-binding;  
 CC KW Transferrase; Serine/threonine-protein kinase; ATP-binding;  
 CC KW Phosphorylation; Magnesium; Membrane.  
 CC FT DOMAIN 158 399 PROTEIN KINASE.  
 CC FT NP\_BIND 164 172 ATP (BY SIMILARITY).  
 CC FT BINDING 185 185 ATP.  
 CC FT ACT\_SITE 269 269 BY SIMILARITY.  
 CC FT DOMAIN 56 62 POLY-GLY.  
 CC FT DOMAIN 668 671 POLY-PRO.  
 CC FT DOMAIN 698 701 POLY-GLU.  
 CC FT DOMAIN 753 758 K->A: NO CATALYTIC ACTIVITY.  
 CC FT MUTAGEN 185 185 E->A: NO CHANGE.  
 CC FT MUTAGEN 192 192 V->A (IN REF. 2).  
 CC FT MUTAGEN 18 18 KL->N (IN REF. 2).  
 CC FT CONFLICT 28 29 S->T (IN REF. 2).  
 CC FT CONFLICT 382 362 EQ->D (IN REF. 2).  
 CC FT CONFLICT 494 495 N->D (IN REF. 2).  
 CC FT CONFLICT 517 517 E->G (IN REF. 2).  
 CC FT CONFLICT 794 794 E->G (IN REF. 2).  
 CC SEQUENCE 888 AA; 96083 MW; CFEFCFD34F889AMB CRC64;  
 CC  
 CC Query Match 23.8%; Score 565.5; DB 1; Length 888;  
 CC Best Local Similarity 37.8%; Freq. No. 1,4e-29;  
 CC Matches 135; Conservative 71; Mismatches 116; Indels 35; Gaps 13;  
 CC  
 CC QY 10 QIKRDLQFPENCGGSGFSGVYRAKWIQDEKAVKLLIKIEEAEI--LSVLSHRNIIQ 67  
 CC DB 152 EEPPEELIDLOWSSGAGVFLGRF--HGEYAVKVRRL-KETDIKRLKHPNIIIT 208  
 CC QY 68 FYGYLEPPNYGVTEVASIGSLVDYNSNRSEMDHIMTATDYAKMHLINAEV 127  
 CC DB 209 FKVGCTGAPCYCIIEMFECAQGLYEVRAGRPVTPSL--LVDSMGJAGMNVLIHL-- 263  
 CC QY 128 KVIRDLKSRNVVIAAGCVKIKICDFGASR-FHNHTTHSLVGFPPMAPEVIOSLPVSET 186  
 CC DB 264 KIHRLDKSPMLITVDVVKISDFTGSKELSDKSTKMSFAGIVAMPAPEVINPEPSEK 323

QY 187 CDTSYGVVLMEMLTREVPFKLEGLOVAMLVKERNRLTIPSSCPSPFAELLHQCNEAD 246  
 DB 324 VDIWSEGVVLMELTRELTPYDVDSALIWGSGNSLHLPVPSCCPDGFKILLROCNWSK 383  
 QY 247 AKRSPFROIISLESMSNDT-SLPDKNSFLNKAERCEIETLERK-----KLE 298  
 DB 364 PNNRSPFQIILLHDIASADVLTPOE--TYFKSOAEWRREVKHAFKIKISEGTCLHLE 441  
 QY 299 RLKSKDEQ-----LKERERLKMEOKLTQSNTPILLPLAARMSSESYFESKTEE 350  
 DB 442 EELVVRERREELHRLDIREH---YERLTERANN--LYMELNALMLQ---LELERE 489  
 CC  
 CC RESULT 5  
 CC M3C\_RAT ID M3C\_RAT STANDARD; PRT; 888 AA.  
 CC AC 063796;  
 CC DT 16-OCT-2001 (Rel. 40, Created)  
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 CC DE Mitogen-activated protein kinase kinase 12 (EC 2.7.1.37)  
 CC GN MAPK12 OR MUK.  
 CC OS Rattus norvegicus (Rat).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 CC OX NCBI\_Taxid=10116;  
 CC RN [1]  
 CC RP MEDLINE=96226099; PubMed=8637721;  
 CC RA Hirai S., Izawa M., Osada S., Sgyrou G., Ohno S.;  
 CC RT MEKK and MUK.";  
 CC RL Oncogene 12:641-650(1996).  
 CC CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.  
 CC Phosphorylates beta-casein, histone 1 and myelin basic protein in  
 CC vitro.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -1- COFACTOR: Magnesium.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (by  
 CC similarity).  
 CC -1- PFM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol  
 CC under basal conditions and dephosphorylated when membrane-  
 CC associated (by similarity).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MAP KINASE KINASE SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; D49785; BA08621.1; -  
 CC DR HSSP; P12931; IFMK.  
 CC DR InterPro; IPR000719; Euk\_pkinase.  
 CC DR InterPro; IPR004040; STY\_pkinase.  
 CC DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 CC DR Pfam; PF00069; Pkinase; 1.  
 CC DR ProDom; PD000001; Euk\_pkinase; 1.  
 CC DR SMART; SM00221; STYKC; 1.  
 CC DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE\_NEG.  
 CC DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 CC DR PROSITE; PS50011; Serine/threonine-protein kinase; ATP-binding;  
 CC KW Transferrase; Serine/threonine-protein kinase; ATP-binding;  
 CC KW Phosphorylation; Magnesium; Membrane.  
 CC FT DOMAIN 158 399 PROTEIN KINASE.  
 CC FT NP\_BIND 164 172 ATP (BY SIMILARITY).  
 CC FT BINDING 185 185 ATP (BY SIMILARITY).  
 CC FT ACT\_SITE 269 269 BY SIMILARITY.  
 CC FT DOMAIN 56 62 POLY-GLY.



[illegible]

```

Db 27 NPEEIDYKELEVEEYVORGAFVVCXAKM--RAKDVAIK---QIESESEKAFIVELRQL 81
Qy 58 SVLSHRENIQFYGVILEPPNYGIIVTEYASIGSYDYINSNRSEEM--DMDHIMTAITDV 114
Db 82 SRVHVHPITVILYGCALNP--VCLMEYAEBSGSLYNVLHG--ASPLPYTTAAHMSWCLOC 137
Qy 115 AKGMHTYILHMAPKVYVHRDLKSRNVVYLAAG-VLKI CDDEG-ASRPNHHTHMSLVGTFPW 172
Db 138 SQGVAVYHSMQPKALHHRDLKPPILLVAGGIYVLCIDFETACDICTHTMNNK--GSAAM 195
Qy 173 MAPEVIOSLPVESECDTYSYGVVLMEMLTREVPFKLEG--IQVAMLVKNERLITPSS 230
Db 196 MAPEVEGNSVNSSECDYFVSGIILMEVITRKPFDEIGCAPRIMW-AVANGTRPPLIKN 254
Qy 231 CPKRSFAELHQCEWADAKKRPSEKQITSL-----ESMSNDT- 267
Db 255 LPKPIESILMTRCWSKDPQSRPSEMEIKYKITHLMRYPGADEPLQYPCQYSDKQSSAT 314
Qy 268 ---SLDPKNSFLNKKEMWCE-LEAVLERLTKLERDLSFEQELKERERRLKMEQKLT 323
Db 315 STGSPMDIASTNINSKSDYMEQVPAINDITIKLESLLKNGAKQSGESGRLLSIGASRG 374
Qy 324 EQSWTPLLPLLAAMESESYFESKTESNSAEMSCQITATSNCEGHGMNPSLOAMLMGF 383
Db 375 SVESLP---PIS-----EGKRMADMSIEARIVATA-GNGQPRRRSIODLTVTGT 421
Qy 384 --GDIFSNMKAGAV 395
Db 422 EPQGVSSRSSSPSV 435

RESULT 8
M3K7 HUMAN STANDARD; PRT; 606 AA.
ID M3K7 HUMAN
AC 04318: O43317; O43319;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase 7 (EC 2.7.1.-)
DE (transforming growth factor-beta-activated kinase 1) (TGF-beta-
DE activated kinase 1).
DE MAPK7 OR TAK1.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Lung;
RA MEDLINE=98153801; PubMed=9480845;
RX Sakurai H., Shigemori N., Haasegawa K., Sugita T.;
RT "TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an
RT NF-kappa B-inducing kinase-independent mechanism.";
RT Biochem. Biophys. Res. Commun. 243:545-549(1998).
CC -1- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKKS.
CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B
CC ACTIVATION.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1A, 1B (SHOWN HERE) AND 1C; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG. TO C210R7.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL; AB009357; BAA25026.1; -.
CC EMBL; AB009356; BAA25025.1; -.

```

DR EMBL: AB009358; BAA25027.2; -  
 DR HSSP: P08631; IADS.  
 DR Genew: HGNC:6859; MAP3K7.  
 DR MIM: 602614; -  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR004040; STY\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR Pfam: PF00069; pkinase; 1.  
 DR ProDom: PD000001; Euk\_pkinase; 1.  
 DR SMART: SM00221; STYK; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR Transferrase; Serine/chreonline-protein kinase; ATP-binding;  
 KW Alternative splicing.  
 FT DOMAIN 8 14  
 FT NP BIND 36 291 POLY-SER.  
 FT BINDING 42 50 ATP (BY SIMILARITY).  
 FT ACT SITE 63 63 ATP (BY SIMILARITY).  
 FT VARSPLIC 156 156 BY SIMILARITY.  
 FT VARSPLIC 404 430 MISSING (IN ISOFORM 1A).  
 FT VARSPLIC 509 518 PLAPCNSKE -> ARTSCRTGPG (IN ISOFORM 1C).  
 FT VARSPLIC 519 606 MISSING (IN ISOFORM 1C).  
 SQ SEQUENCE 606 AA; 67196 MW; 3DBF8147CD174013 CRC64;

Query Match 19.9%; Score 471.5; DB 1; Length 606;  
 Best Local Similarity 29.7%; Pred. No. 1.2e-23;  
 Matches 134; Conservative 82; Mismatches 150; Indels 85; Gaps 20;

QY 7 SFVQIKFDLDFPENCNGSGFSGYRAKMIQDKEVAVKLIKIKKE-----IL 57  
 DB 27 NEEIDYKEIEVEYVGKAFVGVCKAK--BAKDAIK--QIESEERKAFVIELROL 81  
 QY 58 SVLSHRNIIQFYGVILEPPTNGIVTEVASLGLDYINSNRSEEM---DMDHMTWATDV 114  
 DB 82 SRVNPNTVVKYIGACINP--VCLVMEVYEGSLVNVLMG--AEPLPYTAHAMSQCQC 137  
 QY 115 AKGMVYLMHEAPVVKIHRDLKSRNVIAADG-VLKICDFFG-ASRPHNTTMSLVGTFPW 172  
 DB 138 SGGVAYLHSMQKALIHDLKPNLLVAGGVTLKICDFFGACDIQTMHNK--GSAAM 195  
 QY 173 MAPEVYQSLPVSECTDTSYGVVLMEMLTREVPFKLEG--LOYAVLVEKNEELTIPSS 230  
 DB 196 MAPEVFEGSNVSEKCDVSKWGLIMEVITRRKPDDEIGGPAFRIMW-AVHNGTTPPLIKN 254  
 QY 231 CPRSAELHOCWEADAKKRSFKQIISL-----ESMSNDT- 267  
 DB 255 LPKPIESLMTKCSKDPQSRPSMEIYKIMTHLMRYFGADEPLQYPOVSDGQNSAT 314  
 QY 268 ---SLPDKCSFLNKKAMRCE-TEATLERKIKLERDLSFEQELKERERLKKMEQKLT 323  
 DB 315 STGSFMDASINTSKSDTNEQVPATNDTKRLSKLKNQAQSESGRLSGASRG 374  
 QY 324 EQSNTPLLLPLAAMSESYFESTESNSAMSCQITATS---NEEGHGMNSLOAM 379  
 DB 375 SVESLP---PFS-----EGKMSADMSIEARLAIATTAISKPKRGKRTAS----- 417  
 QY 380 LMGEGDIFSMMKAGAVHSGMQIMQAKNS 410  
 DB 418 ---FGNILDVPE---IVISG---NGOPRRRS 439

RESULT 9  
 ABL\_MLVAB  
 ID ABL\_MLVAB STANDARD; PRT; 746 AA.  
 AC P00521;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUN-2002 (Rel. 01, Last annotation update)  
 DE Tyrosine-protein kinase transforming protein ABL (EC 2.7.1.112).  
 GN V-ABL.  
 OS Abelson murine leukemia virus.  
 VC Viruses; Retrovirdae; Retroviridae; Mammalian type C retroviruses.

OX NCBI\_TaxID=11788;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83221648; PubMed=6304726;  
 RA Reddy E.P., Smith M.J., Srinivasan A.;  
 RT "Nucleotide sequence of Abelson murine leukemia virus genome:  
 RT structural similarity of its transforming gene product to other onc  
 RT gene products with tyrosine-specific kinase activity.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:3623-3627(1983).  
 RN  
 RP REVISIONS TO 588-746.  
 RA Reddy E.P., Smith M.J., Srinivasan A.;  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:7372-7372(1983).  
 RN  
 RP SEQUENCE OF 233-327 FROM N.A.  
 RX MEDLINE=83245023; PubMed=6191223;  
 RA Groffen J., Heisterkamp N., Reynolds F.H. Jr., Stephenson J.R.;  
 RT "Homology between phosphotyrosine acceptor site of human c-abl and  
 RT viral oncogene products";  
 RL Nature 304:167-169(1983).  
 CC  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-ABL  
 CC POLYPEPTIDE.  
 CC  
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL  
 CC SUBFAMILY.  
 CC  
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: V01541; -; NOT ANNOTATED\_CDS.  
 DR EMBL: X00010; AAA46470.1; -  
 DR PIR: A00627; TYMVGK.  
 DR HSSP: P00519; 2ABL.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR000980; SH2.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00017; SH2; 1.  
 DR Pfam: PF00069; pkinase; 1.  
 DR PRINTS: PR00401; SH2DOMAIN.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR ProDom: PD000001; Euk\_pkinase; 1.  
 DR ProDom: PD000093; SH2; 1.  
 DR SMART: SM00252; SH2; 1.  
 DR SMART: SM00219; TyKc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Polyprotein; Tyrosine-protein kinase; Transferase; Oncogene;  
 KW SH2 domain.  
 FT DOMAIN 13 103  
 FT NP BIND 128 379 SH2.  
 FT BINDING 128 379 PROTEIN\_KINASE.  
 SQ SEQUENCE 746 AA; 81872 MW; B9072FF5F5FE9257 CRC64;

Query Match 18.3%; Score 434; DB 1; Length 746;  
 Best Local Similarity 28.7%; Pred. No. 4.4e-21;  
 Matches 121; Conservative 73; Mismatches 172; Indels 56; Gaps 13;  
 QY 15 DLQFFENCNGSGFSGYRAKMIQDKEVAVKLL-----LKIE--KEAELSLSHRNIIQ 67  
 DB 127 DITMKHKLGGGQGVGVVGVKVSLLFAVATLKEDTVEEFLKEAAVMEKIHGPNLVQ 186  
 QY 68 FYGVILEPPTNGIVTEVASLGLDYINSNRSEEMDMDHMTWATDVAKGMVYLMHEAPV 127  
 DB 187 LLGVCTRRPPFYIITTEPNTYGNLDYLRCKRQEVSAVLLVMATQISSAMEYLEKK--- 243

QY 128 KVIHRLKSRNVYIADGVKICDFGASRPHNTHMVLGT-FP--WMAPEVIOSLPVS 184  
 DB 244 NFTHRLAANNCVGNHLYKAVDFGSLRMTGDTYAHGAKFPIKMTAPESIAVYKFS 303  
 QY 185 ETCDTYSYGVLMEMLTREV-PFKGLGLOVAMLVKERNLTPSSCCPSFALHQCW 243  
 DB 304 IKSDVMAFVGLTWELATYAGMSPPYGDIDSGV-YELLEKDYRMERPECCPKRYELMAGW 362  
 QY 244 EADAKRPPSPFOIISLESMSNDTSLPDKNSFLHNKAEMRCIEATLETLKLEKD-- 300  
 DB 363 QWNSDPSPEFHQAFETMFQESSISD-----EVEKELG-KRGTRGAG 406  
 QY 301 --LSFKEQLKERERRLKMEQKLTQESNTPLLPLAARMSSESYFESKTESNSAEMSC 358  
 DB 407 SMLQAPLPTTRTCR-RAAQKASPSPLTFPKLRRQVTASPSGSLSHKKEATKGSASGM 465  
 QY 359 QITATSNGEHGMPSLOAMMLMGFDIFSNNKAGAVHSGMQINMOAKONSSKTTSKR 418  
 DB 466 GTPATA---EPAPPS-----NKVGLSKASSEEMRVRRKHSSSPGRDK 506  
 QY 419 GK 420  
 DB 507 GR 508

RESULT 10  
 ID ABL1 HUMAN STANDARD; PRT; 1130 AA.  
 AC P00519; 016133; 013869; 013870;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Proto-oncogene tyrosine-protein kinase ABL1 (EC 2.7.1.112) (p150)  
 DE (c-ABL).  
 GN ABL1 OR ABL OR JTK7.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 OX 11  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fibroblast;  
 RX MEDLINE=90082420; PubMed=2587768;  
 RA Fainstein E., Elnat E., Gokkel E., Marcelle C., Croce C.M.,  
 RA Gale R.P., Canani E.;  
 RT "Nucleotide sequence analysis of human abl and bcr-abl cDNAs.";  
 RL Oncogene 4:1477-1481(1989).  
 RN 12  
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.  
 RX MEDLINE=87028219; PubMed=3021337;  
 RA Shliveiman E., Lifshitz B., Gale R.P., Roe B.A., Canani E.;  
 RT "Alternative splicing of RNAs transcribed from the human abl gene and  
 RT from the bcr-abl fused gene.";  
 RL Cell 47:277-284(1986).  
 RN 13  
 RP SEQUENCE FROM N.A. (ISOFORMS IA AND IB).  
 RC TISSUE=Lung carcinoma;  
 RX MEDLINE=95394474; PubMed=7665185;  
 RA Chisoe S.L., Bodenteich A., Wang Y.-F., Wang Y.-P., Burian D.,  
 RA Clifton S.W., Crabtree J., Freeman A., Iyer K., Jian L., Ma Y.,  
 RA McLaurin H.-J., Pan H.-Q., Sathan O.H., Toch S., Wang Z., Zhang G.,  
 RA Heisterkamp N., Groffen J., Roe B.A.;  
 RT "Sequence and analysis of the human ABL gene, the BCR gene, and  
 RT regions involved in the Philadelphia chromosome translocation.";  
 RL Genomics 27:67-82(1995).  
 RN 14  
 RP SEQUENCE OF 360-426 FROM N.A.  
 RX MEDLINE=83245023; PubMed=6191223;  
 RA Groffen J., Heisterkamp N., Reynolds F.H. Jr., Stephenson J.R.;  
 RT "Homology between phosphotyrosine acceptor site of human c-abl and  
 RT viral oncogene products.";  
 RL Nature 304:167-169(1983).  
 RN 15

RP SEQUENCE OF 27-40 FROM N.A.  
 RX MEDLINE=86065859; PubMed=2825022;  
 RA Fainstein E., Marcelle C., Rosner A., Canani E., Gale R.P.,  
 RA Drazzen O., Smith S.D., Croce C.M.;  
 RT "A new fused transcript in Philadelphia chromosome positive acute  
 RT lymphocytic leukaemia.";  
 RL Nature 330:386-388(1987).  
 RN 16  
 RP SEQUENCE OF 825-845 FROM N.A.  
 RX MEDLINE=9414231; PubMed=7545908;  
 RA Inokuchi K., Futaki M., Dan K., Nomura T.;  
 RT "Sequence analysis of the mutation at codon 834 and the sequence  
 RT variation of codon 837 of c-abl gene.";  
 RL Leukemia 8:343-344(1994).  
 RN 17  
 RP STRUCTURE BY NMR OF SH2 DOMAIN.  
 RX MEDLINE=92370689; PubMed=1505033;  
 RA Overduin M., Rios C.B., Mayer B.J., Baltimore D., Cowburn D.;  
 RT "Three-dimensional solution structure of the src homology 2 domain of  
 RT c-abl.";  
 RL Cell 70:697-704(1992).  
 RN 18  
 RP STRUCTURE BY NMR OF SH2 DOMAIN.  
 RX MEDLINE=93101588; PubMed=1281542;  
 RA Overduin M., Mayer B.J., Rios C.B., Baltimore D., Cowburn D.;  
 RT "Secondary structure of Src homology 2 domain of c-Abl by  
 RT heteronuclear NMR spectroscopy in solution.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:11673-11677(1992).  
 RN 19  
 RP STRUCTURE BY NMR OF SH3 DOMAIN.  
 RX MEDLINE=96131878; PubMed=8590002;  
 RA Gosser Y.Q., Zheng J., Overduin M., Mayer B.J., Cowburn D.;  
 RT "The solution structure of Abl SH3, and its relationship to SH2 in  
 RT the SH(32) construct.";  
 RL Structure 3:1075-1086(1995).  
 RN 20  
 RP 3D-STRUCTURE MODELING OF SH3 DOMAIN.  
 RX MEDLINE=9519229; PubMed=7892170;  
 RA Pisabarro M.T., Ortiz A.R., Serrano L., Wade R.C.;  
 RT "Homology modeling of the Abl-SH3 domain.";  
 RL Proteins 20:203-215(1994).  
 RN 21  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 57-218.  
 RX MEDLINE=96398698; PubMed=8805596;  
 RA Nam H.-J., Hasser W.G., Roberts T.M., Frederick C.A.;  
 RT "Intramolecular interactions of the regulatory domains of the Bcr-Abl  
 RT kinase reveal a novel control mechanism.";  
 RL Structure 4:1105-1114(1996).  
 RN 22  
 RP X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS) OF 64-121.  
 RX MEDLINE=98365516; PubMed=9698566;  
 RA Pisabarro M.T., Serrano L., Wilmanns M.;  
 RT "Crystal structure of the Abl-SH3 domain complexed with a designed  
 RT high-affinity peptide ligand: implications for SH3-ligand  
 RT interactions.";  
 RL J. Mol. Biol. 281:513-521(1998).  
 RN 23  
 RP -1 CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1 SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1 ALTERNATIVE PRODUCTS: 2 ISOFORMS; IA (SHOWN HERE) AND IB; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1 TISSUE SPECIFICITY: WIDELY EXPRESSED.  
 CC -1 DISEASE: PARTICIPATES IN A T(9;22)(Q34;Q11) CHROMOSOMAL  
 CC TRANSLOCATION THAT PRODUCES A BCR-ABL ONCOGENE RESPONSIBLE FOR  
 CC CHRONIC MYELOID LEUKEMIA (CML), ACUTE MYELOID LEUKEMIA (AML), AND  
 CC ACUTE LYMPHOBLASTIC LEUKEMIA (ALL).  
 CC -1 SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL  
 CC SUBFAMILY.  
 CC -1 SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
 CC -1 SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC -1 DATABASE: NME=Atlas Genet. Cytogenet. Oncol. Haematol.;  
 CC WWW="http://www.infobio.gen.fr/services/chromocancer/genes/ABL.html".  
 CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X16416; CAA34438.1; -  
 DR EMBL: M14752; AAS15161.1; -  
 DR EMBL: U07563; AAB60394.1; -  
 DR EMBL: U07563; AAB60393.1; -  
 DR EMBL: U07561; AAB60393.1; JOINED.  
 DR EMBL: S69223; AAD14034.1; -  
 DR PIR: A25582; TYHUA.  
 DR PDB: 1AB2; 31-JAN-94.  
 DR PDB: 1AB2; 01-NOV-94.  
 DR PDB: 2AB1; 04-SEP-97.  
 DR PDB: 1AMC; 28-JAN-98.  
 DR PDB: 1BB2; 25-NOV-98.  
 DR GeneW: HGNC:76; ABL1.  
 DR MIM: 189980; -  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR000980; SH2.  
 DR InterPro: IPR01452; SH3.  
 DR InterPro: IPR01245; Tyr\_pkinase.  
 DR Pfam: PF00017; SH2; 1.  
 DR Pfam: PF00018; SH3; 1.  
 DR Pfam: PF00069; pkinase; 1.  
 DR PRINTS: PR00401; SH2DOMAIN.  
 DR PRINTS: PR01019; TYRKINASE.  
 DR ProDom: PD000001; Euk\_pkinase; 1.  
 DR ProDom: PD000066; SH3; 1.  
 DR ProDom: PD000093; SH2; 1.  
 DR SMART: SM00252; SH2; 1.  
 DR SMART: SM00326; SH3; 1.  
 DR SMART: SM00219; TyKc; 1.  
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE: PS00109; PROTEIN KINASE TYR; 1.  
 DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00002; SH2; 1.  
 DR PROSITE: PS00002; SH3; 1.  
 DR Transferrase; Tyrosine-protein kinase; Proto-oncogene; ATP-binding;  
 DR Phosphorylation; SH2 domain; SH3 domain; Chromosomal translocation;  
 DR 3D-structure; Alternative splicing.  
 DR DOMAIN 61 121  
 FT DOMAIN 127 217  
 FT DOMAIN 242 493  
 FT DOMAIN 605 609  
 FT DOMAIN 18 22  
 FT DOMAIN 605 609  
 FT DOMAIN 782 1019  
 FT DOMAIN 897 903  
 FT SITE 26 27  
 FT NP\_BIND 248 256  
 FT BINDING 271 271  
 FT ACT\_SITE 363 363  
 FT MOD\_RES 393 393  
 FT VARSPIC 1 26  
 FT CONFLICT 140 140  
 FT CONFLICT 159 159  
 FT CONFLICT 424 425  
 FT CONFLICT 445 445  
 FT CONFLICT 459 459  
 FT CONFLICT 520 520  
 FT CONFLICT 719 719  
 FT CONFLICT 837 837  
 FT CONFLICT 837 837  
 FT CONFLICT 863 863  
 FT CONFLICT 894 894

FT CONFLICT 917 919 SPS -> RPG (IN REF. 2).  
 FT CONFLICT 952 952 G -> A (IN REF. 2).  
 FT CONFLICT 967 968 OS -> HP (IN REF. 2).  
 FT CONFLICT 983 983 S -> LS (IN REF. 2).  
 FT CONFLICT 1022 1022 MISSING (IN REF. 2).  
 FT CONFLICT 1045 1045 R -> G (IN REF. 2).  
 FT CONFLICT 1103 1103 T -> S (IN REF. 2).  
 FT TURN 120 120

Query Match 17.3%; Score 410; DB 1; Length 1130;  
 Best Local Similarity 25.9%; Pred. No. 2,7e-19;  
 Matches 119; Conservative 80; Mismatches 183; Indels 78; Gaps 15;

QY 15 DLQFENGCGSSFGSVYRAKISODKEVAYKLTKEEKEAEITLSVSHRNIIQ 67  
 DB 241 DITMKHKLGGQYGEVYEGVKKYSLTVAAYKLTEDMEVEEFLKEAAVVKKEIGHPIVQ 300

QY 68 FYGVILEPPNYGIYTEYASLGSLDYINSNRSEMDHIMTATDVAKGMHYHMEAPV 127  
 DB 301 LLGVCTREPPPIYITTEMTGNLLDYLRCCNRQEVAVVILYMAQTSSAMEYLEKK-- 357

QY 128 KVTHRDLSKRVVYAADGVLTICDFGASRPHNHTHNSIVGT-PP-WMAPEVIQSLPV 184  
 DB 358 NFTHRDLAARNCLVGENHLYKVADFGISRLMTGDTYTAHAGAKFPIKWTAPESLAVNKF 417

QY 185 ETCDTSYGVVLMEMLTREV-PPKGLGLOVAVLVYKERNLTIIPSSCPSPFALLHQCW 243  
 DB 418 IKSDVMAFVULMEIATYGMSPYFGIDLSOV-YELLEKDKRMKRPFGCEPKVVELRWAC 476

QY 244 EADAKKRPSPKQIISLESNSNDTSLPDKNSFLHNKAVERCEIATELTLEKLERDLSF 303  
 DB 477 QWNSDRPSFAEIHQAFETFEQSSISDEVEKELGQG-----VRNAVSTL-----LQA 525

QY 304 KEDELKERERLKMWFQKLTQCSNTPLLPLAARMBESYFESKTE-ESNSAE---MSC 358  
 DB 526 PELPTKTRTSRAAERHDTTVPEMP-----HSKGGESDPLDHEPAVSP 570

QY 359 QITATNGEGHGNPQLQAMLMFGFDIFSPYKAGAVHMSGQINN-----QAKNSSTK 413  
 DB 571 LIPRKERGPPEG-----GLNEDERLPRKDKTNPLSALIKKKXTAPT 613

QY 414 TSKRRGKYNMALGFSDPLDSEGDHDDDDGEEDNDMN 453  
 DB 614 PPKRS-----SSFREMDOGQERRAGAGEEGRDIN 643

RESULT 11  
 ID ABL1\_MOUSE STANDARD; PRT; 1123 AA.  
 AC P00520; Q61252; Q61253; Q61254; Q61255; Q61256; Q61257; Q61258;  
 AC Q61259; Q61260; Q61261; P97896;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Proto-oncogene tyrosine-protein kinase ABL1 (EC 2.7.1.112) (p150)  
 DE (c-ABL).  
 GN ABL1 OR ABL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=88068561; PubMed=3117402;  
 RA Opti C., Shore S.K., Reddy E.P.;  
 RT "Nucleotide sequence of testis-derived c-abl cDNAs: implications for  
 RL testis-specific transcription and abl oncogene activation.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8200-8204(1987).  
 RN [2]  
 RP SEQUENCE OF 1-187 FROM N.A. (ISOFORMS I TO IV).  
 RX MEDLINE=95394474; PubMed=765185;  
 RA Chisoe S.L., Bodenteich A., Wang Y.-F., Wang Y.-P., Burian D.,

DR	pfam; PF00017; SH2; 1.
DR	pfam; PF00018; SH3; 1.
DR	pfam; PF00069; pkinase; 1.
DR	PRINTS; PR00401; SH2DOMAIN.
DR	PRINTS; PR00109; TYRKINASE.
DR	Prodom; PD000001; Euk_pkinase; 1.
DR	Prodom; PD000066; SH3; 1.
DR	Prodom; PD000093; SH2; 1.
DR	PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR	PROSITE; PS00109; PROTEIN KINASE TYR; 1.
DR	PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE; PS50001; SH2; 1.
KW	Transferase; Tyrosine-Protein kinase; Proto-oncogene; ATP-binding;
KW	Phosphorylation; SH2 domain; SH3 domain; Alternative splicing;
KM	Nuclear protein; 3D-structure.
FT	DOMAIN 61 121 SH3.
FT	DOMAIN 127 217 SH2.
FT	DOMAIN 242 493 PROTEIN KINASE.
FT	DOMAIN 605 609 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT	DOMAIN 18 22 POLY-SER.
FT	DOMAIN 605 609 POLY-LYS.
FT	DOMAIN 804 1012 PRO-RICH.
FT	NP_BIND 891 897 POLY-PRO.
FT	ACT_SITE 248 256 ATP (BY SIMILARITY).
FT	BINDING 271 271 ATP (BY SIMILARITY).
FT	MOD_RES 353 363 BY SIMILARITY.
FT	VARSPLIC 1 26 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	VARSPLIC 1 26 MLEICLKLVGCKSKGLSSSSCYCLE -> MSQRTITKCR
FT	VARSPLIC 1 26 VORDPALFPM (IN ISOFORM II).
FT	VARSPLIC 1 26 MLEICTKLTVGCKSKGLSSSSCYCLE -> MISFDLSDEL
FT	VARSPLIC 1 26 HLKILIVLV (IN ISOFORM II).
FT	VARSPLIC 1 26 MLEICTKLTVGCKSKGLSSSSCYCLE -> MGQOPGYVLGD
FT	VARSPLIC 1 26 ORREPLRLHPITKGAKGRDSSRHGGPHCNVVEH (IN
FT	VARSPLIC 1 26 ISOFORM IV).
FT	CONFLICT 184 187 LYS -> VGMV (IN REF. 2).
FT	SEQUENCE 1123 AA; 122676 MW; 284RF030644AFDF CRC64;
Query Match	17.1%; Score 405; DB 1; Length 1123;
Best Local Similarity	27.2%; Pred. No. 5.7e-19;
Matches 125; Conservative	78; Mismatches 179; Indels 78; Gaps 17,
DQ	15 DLDPFNCGGSGSYRAKMISODKEVAVVKL---LKIE--KEAELIIVLSHRNIIQ 67
DQ	241 DIRMKHLTGSGGQGEYEGVWKYSLTVAVKTLKEDTMVEEELKEAAVMKEIKHPNLVQ 300
DQ	68 FYGVILEPPNYIGVTVEYASLGSIYDTINSRSEMDHIIMATPVAKGHYLHEAPV 127
DQ	301 LLGCTEEPEPYIITEFMYIGNLLDIARECNRQESAVVLLYATQISAMEYLEKK--- 357
DQ	128 KVTHRDLSKNVVIAADGVLTICDPFGASREFRNHTTMSIVGT-FP--WMAPEVIQSLPVS 184
DQ	358 NFTHRDLAARNCLVGENHLVYKADFGLSRMTGDDITTAHGAKEPIKTAPESIALVNKES 417
DQ	185 ETCDTSYGVVLMEMLTREV-PFKGLEGLOVAMLYVEKERLTIIPSSCPSPFAELHQCV 243
DQ	418 IKSDVMAFGVILWEIATYAGMSYPYGIDLSQV-YELLEKDVRMERPGCEPKAYVELMPACW 476
DQ	244 EADAKRPSPFOIISIESMSNDTSLPDCKNSFLHNKAEMRCIEATLETKLIERD--- 300
DQ	477 QWNPSDRSPFAELHQAETMFQESSISD-----EVEKELG-KKGTREGAG 520
DQ	301 --LSFKEQLKERERLRKMMEOUKTEOSTNPILLPLAARMSBSYFSFKTESNSAEMSC 358
DQ	521 SMLQAPELPTTRICR-RAABOK-DAPDPELL-----HYKGLGSBDALDSEPA-VSP 570
DQ	359 QITATNSGEHGQNPISLOAMLMGFQITFSNKRGAAVHSGMQNM-----QAKONSST 413
DQ	571 ILPRKERGPDPG-----SLNEBERLLPRDKKNLNFSAIIKKKKKAAAPT 613
DQ	414 TSKRRGKKVMNALGPSFDLSSEGGDDDDDDDEEDENDMDN 453
DQ	614 PPKRS-----SSFREMDQGPDRGASEDDDSRELNC 643



RESULT 12  
 ABL2\_HUMAN STANDARD; PRT; 1182 AA.  
 AC P42684;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tyrosine-protein kinase ABL2 (EC 2.7.1.112) (Tyrosine kinase ARG).  
 GN ABL2 OR ARG OR ABL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=90332670; PubMed=2198571;  
 RT "Kruh G.D., Perego R., Miki T., Aaronson S.A.;  
 RT "The complete coding sequence of arg defines the Abelson subfamily of  
 RT cytoplasmic tyrosine kinases.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:5802-5806(1990).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1A AND 1B (SHOWN HERE); ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.  
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL  
 CC SUPERFAMILY.  
 CC -1- SIMILARITY: CONTRAINS 1 SH2 DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC -1- DATABASE: NAME=AbLac Genet. CytoGenet. Oncol. Haematol.;  
 CC WWW="http://www.infobiogen.fr/services/chromancer/Gene/ABL2ID226.html".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; M35296; AAA3553.1; -.  
 DR HSSP; P00519; 1BBZ.  
 DR Genew; HGNC:77; ABL2.  
 DR MIM; 164690; -.  
 DR InterPro; IPR000719; Euk\_Pkinase.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001245; Tyr\_Pkinase.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR PRINTS; PRO0401; SH2DOMAIN.  
 DR PRINTS; PRO0109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_Pkinase; 1.  
 DR ProDom; PD000065; SH3; 1.  
 DR ProDom; PD000093; SH2; 1.  
 DR SMART; SM00252; SH2; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR SMART; SM00219; TYRKC; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS50001; SH2; 1.  
 DR PROSITE; PS50002; SH3; 1.  
 DR TransErase; Tyrosine-protein kinase; Proto-oncogene; ATP-binding;  
 KW Phosphorylation; SH2 domain; SH3 domain; Alternative splicing.  
 FT DOMAIN 173 167  
 FT DOMAIN 173 263  
 FT DOMAIN 288 539  
 FT DOMAIN 561 564  
 RP POLY-SER.  
 RP PROTEIN KINASE.

FT DOMAIN 658 660 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT DOMAIN 732 739 POLY-GLY.  
 FT DOMAIN 843 1055 PRO-RICH.  
 FT DOMAIN 984 988 POLY-PRO.  
 FT NP\_BIND 294 302 ATP (BY SIMILARITY).  
 FT BINDING 317 317 ATP (BY SIMILARITY).  
 FT ACT\_SITE 409 409 BY SIMILARITY.  
 FT MOD\_RES 439 439 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT VAAPSLIC 1 73 MGQGVGVGAPGGLQDPGIRGSSARSGRRDPAGRT  
 FT VLLPNTYGRDQDTSLCCLCTEASESALPDLT (IN  
 FT ISOFORM 1A).  
 SQ SEQUENCE 1182 AA; 128343 MM; ED93869BCB14FAA CRC64;  
 Query Match 17.1%; Score 405; DB 1; Length 1182;  
 Best Local Similarity 32.8%; Pred. No. 6.1e-19;  
 Matches 113; Conservative 59; Mismatches 144; Indels 28; Gaps 13;  
 QY 5 GASFWQIKFD----DLQFENCGGSGFSGVYRAKWIQDKEVAVKXU-----LKIE---KE 53  
 DB 273 GVSPIHDKWEMERTDIMKHTGSGGYGEYVGVKXKSLTVAVKTLKEDTMEVEEFKE 332  
 QY 54 AEIISVLSHNNITQFYGV-ILPPNNGIVTEYASLSGLYDINSNRSEEMDMHIMTWAT 112  
 DB 333 AAVKKEIKHPVLVQLGCTLEPPFY-IVTEYMPYGNLLDYLRCKNREEVTAVALVLMAT 391  
 QY 113 DVAKGMHYLMEAPVKYIHVDLKSARNVVIADGVLKICDYGASFFNHTTMSLVGT-PP 171  
 DB 392 QISSAMELEKK--NFIHDLAARNCLVGENHVKADGSLMTMDTYTNAAGAKFP 448  
 QY 172 --WNAPEVIOQLPYSETCDTYSYGVVLMELTREV--PFKGLGGLQVWLVEKNERLTIP 228  
 DB 449 IKWTAPESLAVNFMFSISDVWAFGVLMETATYGMSPYPIDLSQV-YDLLEKGYRMEQP 507  
 QY 229 SSCRSRAELHQCEADAKRPSFKQIISILSMSNDTSLPDKNSFLNKAEMRCEIE 288  
 DB 508 EGCEPPKYLELMRAKWKXSPADRSFAETHQAFETMFHDSISSEVAEL-GRASSSSVY 566  
 QY 289 ATLERLKLEKRLDSFKQELKERERLKWME--QKLTQENTPL 330  
 DB 567 PYLRLPLIPS-----KTRILKOVENKENTEGADTENSASSL 606  
 RESULT 13  
 ABL\_DROME STANDARD; PRT; 1520 AA.  
 ID ABL\_DROME  
 AC P00522;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-JUN-1990 (Rel. 13, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tyrosine-protein kinase Abl (EC 2.7.1.112) (D-ash).  
 GN ABL OR DASH OR ABL-1.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RP MEDLINE=86174728; PubMed=2832740;  
 RA Henkemeyer M.O., Bennett R.L., Gertler F.B., Hoffmann F.M.;  
 RT "DNA sequence, structure, and tyrosine kinase activity of the  
 RT Drosophila melanogaster Abelson proto-oncogene homolog.";  
 RL Mol. Cell. Biol. 8:843-853(1988).  
 RN (2)  
 RP SEQUENCE OF 374-648 FROM N.A.  
 RX MEDLINE=84082064; PubMed=6317185;  
 RA Hoffmann F.M., Fresno L.D., Hoffman-Falk H., Shilo B.-Z.;  
 RT "Nucleotide sequences of the Drosophila src and abl homologs:  
 RT conservation and variability in the src family oncogenes.";  
 RL Cell 35:393-401(1983).  
 RN (3)  
 RP FUNCTION.

RX MEDLINE=9829828; PubMed=9635189;  
 RA Loureiro J., Peifer M.;  
 CC "Role of Armadillo, a Drosophila catenin, during central nervous  
 RT system development.";  
 RL Curr. Biol. 8:622-632(1998).  
 CC -1- FUNCTION: ARM AND ABL PROTEINS FUNCTION COOPERATIVELY AT ADHERENS  
 CC -1- JUNCTIONS IN BOTH THE CNS AND EPIDERMIS.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL  
 CC SUBFAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL, M13692; AAA28934.1; -  
 DR EMBL, M13690; AAA28934.1; JOINED.  
 DR EMBL, M13691; AAA28934.1; JOINED.  
 DR EMBL, K01042; AAA28443.1; -  
 DR PIR, A28128; TVFEFA.  
 DR HSSP, P00519; IAB2.  
 DR FlyBase; FBpr0000017; Abl.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00401; SH2DOMAIN.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR ProDom; PD000066; SH3; 1.  
 DR ProDom; PD000093; SH2; 1.  
 DR SMART; SM00252; SH2; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR SMART; SM00219; Tyrc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS50001; SH2; 1.  
 DR PROSITE; PS50002; SH3; 1.  
 DR Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;  
 KM SH2 domain; SH3 domain.  
 FT DOMAIN 204 265  
 FT DOMAIN 271 363  
 FT DOMAIN 388 644  
 FT NP\_BIND 394 402  
 FT BINDING 417 417  
 FT ACT\_SITE 509 509  
 FT MOD\_RES 539 539  
 FT CONFLICT 374 377  
 FT CONFLICT 645 648  
 FT SEQUENCE 1520 AA; 161836 MW; AD6A5060579FAD7B CRC64;

Query Match 16.9%; Score 402; DB 1; Length 1520;  
 Best Local Similarity 33.8%; Pred. No. 1,3e-18;  
 Matches 94; Conservative 50; Mismatches 119; Indels 15; Gaps 7;

QY 15 DLQFFNCGGSGFSGYVRAKMSIDQKEVAVKCL-----LK-TEKEAEILSVLSHRNIQ 67  
 DB 387 DLMKRLGGGQGYEYEAIVKGNVAVKTLKEPTMLKDLLEAAIWKKEKHPNIVQ 446  
 QY 68 FYGVILEPFPYGVITEVAGSLVYDINSRSEMDMIMTATVDVAKAMHILNHAAPV 127

DB 447 LIGVCTREPPFYITTEFMSHGNLIDFLNSAGRETLDAVALYMTAQIAGMSYLESR--- 503  
 QY 128 KYIHRDLKSRNVVIAADGVLKICDGFASRHHHTHMSLVGT-PP--WMAPEVIOQLPV 184  
 DB 504 NYIHRDLARNCLVGNKLVKVAADPGLARMDDTYTAAGAKFPKIKTAPBGLAVNKS 563  
 QY 185 EFCDTYSYGVVIMEMLTREV-PFKGLEGLQVAMLVYKNERLTITSSCPSPFAELLHQCV 243  
 DB 564 TKSVDVAPFVLMEIATYGMSPYPAID-LTDVYHKLDKGYRMRPBGCPREYVDLMRCQW 622  
 QY 244 EADAKRPSFKQIISTLESMSNDTSLPKCNSFLNKA 281  
 DB 623 QMDATDRPTFKSIHNLBEMFOESSITTEAVEVKQANNA 660

RESULT 14  
 ID KYKI D1CDI STANDARD; PRT; 1584 AA.  
 AC P18160;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Non-receptor tyrosine kinase spore lysis A (EC 2.7.1.112) (Tyrosine-  
 DE protein kinase 1).  
 GN PYKA OR SPYA OR DEYK1.  
 OS Dictyostelium discoideum (slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
 OX NCBI\_TaxId=4689;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DH10;  
 RX MEDLINE=97053827; PubMed=8898241;  
 RA Nickolls G.H., Osherev N., Loomis W.F., Spudich J.A.;  
 RT "The Dictyostelium dual-specificity kinase SPYA is essential for  
 RT spore differentiation.";  
 RL Development 122:3285-3305(1996).  
 RN [2]  
 RP SEQUENCE OF 1248-1584 FROM N.A.  
 RX MEDLINE=90287147; PubMed=1972546;  
 RA Tan J.L., Spudich J.A.;  
 RT "Developmentally regulated protein-tyrosine kinase genes in  
 RT Dictyostelium discoideum.";  
 RL Mol. Cell. Biol. 10:3578-3583(1990).  
 CC -1- FUNCTION: ESSENTIAL FOR SPORE DIFFERENTIATION.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT WITH A PEAK  
 CC DURING THE MOUND STAGE OF MORPHOGENESIS.  
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.  
 CC -1- SIMILARITY: CONTAINS 1 SAM DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U32174; AAB41125.1; -  
 DR EMBL, M33785; AAA33202.1; -  
 DR PIR; A35670; A35670.  
 DR DictyDB; DD03010; PYKA.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001660; SAM.  
 DR InterPro; IPR003878; SPY domain.  
 DR InterPro; IPR003877; SPY\_receptor.  
 DR InterPro; IPR004040; STY\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00536; SAM; 1.  
 DR Pfam; PF00622; SPY; 3.  
 DR ProDom; PD000001; Euk\_pkinase; 1.

DR SMART; SM00454; SAM; 1.  
 DR SMART; SM00449; SPY; 3.  
 DR SMART; SM00221; STYK; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00105; SAM\_DOMAIN; 1.  
 KM Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.  
 FT DOMAIN 908 972 SAM.  
 FT DOMAIN 403 420 POLY-ASN.  
 FT DOMAIN 428 435 POLY-THR.  
 FT DOMAIN 449 480 POLY-ASN.  
 FT DOMAIN 483 491 POLY-ASN.  
 FT DOMAIN 494 508 POLY-ASN.  
 FT DOMAIN 512 532 POLY-ASN.  
 FT DOMAIN 596 600 POLY-ASN.  
 FT DOMAIN 808 811 POLY-PHE.  
 FT DOMAIN 1026 1029 POLY-SER.  
 FT DOMAIN 1195 1210 POLY-ASN.  
 FT DOMAIN 1215 1220 POLY-GLN.  
 FT DOMAIN 1224 1233 POLY-GLN.  
 FT DOMAIN 1266 1274 POLY-PRO.  
 FT DOMAIN 1289 1561 PROTEIN\_KINASE.  
 FT NP\_BIND 1295 1303 ATP (BY SIMILARITY).  
 FT BINDING 1316 1316 ATP (BY SIMILARITY).  
 FT ACT\_SITE 1417 1417 BY SIMILARITY.  
 FT CONFLICT 1248 1248 D -> R (IN REF. 2).  
 FT CONFLICT 1435 1435 V -> L (IN REF. 2).  
 SQ SEQUENCE 1584 AA; 174304 MW; 5D158945BD8E01E3 CRC64;

Query Match 16.9%; Score 401.5; DB 1; Length 1584;  
 Best Local Similarity 34.6%; Pred. No. 1.5e-18;  
 Matches 102; Conservative 52; Mismatches 104; Indels 37; Gaps 8;

QY 10 QIKPDLOFFENCNGSGSYVRKMIQDKEVAVK-----KLTKKEAEILS 58  
 DB 1283 EIDFNELEFGTIGKFGFGEVKGW--REIDVAKIYRDOFKTKSLVNFQNEVGILS 1340  
 QY 59 VLSHRNITQFPGVILE--PNYGIETVYASLGSLVDYINSNRSEMDMH-MTWATDVA 115  
 DB 1341 KLRHNVVQFLGACTAGGEDHCHITVEMWGGSLNQFLTDHFNLEQNPRIKLALDIA 1400  
 QY 116 KGMVYLHEAVNVYIHRDLKSRNVI-----AADGVLCIKDFGASRFNHTT 162  
 DB 1401 KGMVYLHEAVNVYIHRDLKSRNVI-----AADGVLCIKDFGASRFNHTT 162  
 QY 163 HM-SLVGTFPMMAPEVYIQLPVSECTDYSYGVVLEMTREVPFKGLEQVAVLVE 220  
 DB 1460 SQMTQSVGCIYMAPEVYKGDNSSEKSDYSGVNLFEELTSDPEQDMKMKVLAHAA 1519  
 QY 221 KNEELTIPSSCPSPFAELHQCMEADAKKPSFKQIISILESM-----SNDTSLP 270  
 DB 1520 ESYRPPPIPLTSSKMKELITQCMDSNPDSRPTFKQIIVHLKEMDQGVSPASVP 1574

RESULT 15  
 ABL\_FSVHY STANDARD; PRT; 439 AA.  
 ID ABL\_FSVHY STANDARD; PRT; 439 AA.  
 AC P10447;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tyrosine-protein kinase transforming protein ABL (EC 2.7.1.112).  
 GN V-ABL.  
 OS Feline sarcoma virus (strain Hardy-Zuckerman 2).  
 OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirinae.  
 OC NCBI\_TaxID=11776;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87141338; Pubmed=3029415;  
 RA Bergold P.J., Blumenthal J.A., D'Andrea E., Snyder H.W. Jr.,  
 RA Lederman L., Silverstone A., Nguyen H., Besmer P.,  
 RT "Nucleic acid sequence and oncogenic properties of the H22 feline

RT sarcoma virus v-abl insert";  
 RL J. Virol. 61:1193-1202(1987).  
 CC 1. CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1 MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-ABL-POL  
 CC POLYPEPTIDE.  
 CC -1 SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL  
 CC SUBFAMILY.  
 CC -1 SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
 CC -1 SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M15805; AAA43042.1; -.  
 DR PIR; A26132; TYMVAB.  
 DR HSP; P00519; 1BBZ.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00401; SH2DOMAIN.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR ProDom; PD000066; SH3; 1.  
 DR ProDom; PD000093; SH2; 1.  
 DR SMART; SM00252; SH2; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR SMART; SM00219; TYKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00002; SH3; 1.  
 DR PROSITE; PS00001; SH2; 1.  
 DR PROSITE; PS00002; SH3; 1.  
 KW Polyprotein; Tyrosine-protein kinase; Transferase; Oncogene;  
 KW SH2 domain; SH3 domain.  
 FT DOMAIN 10 70 SH3.  
 FT DOMAIN 76 166 SH2.  
 FT DOMAIN 191 439 PROTEIN\_KINASE.  
 SQ SEQUENCE 439 AA; 50004 MW; 13579BDFED1481AB CRC64;

Query Match 16.2%; Score 385.5; DB 1; Length 439;  
 Best Local Similarity 33.2%; Pred. No. 3.3e-18;  
 Matches 84; Conservative 53; Mismatches 101; Indels 15; Gaps 7;

QY 15 DLOFFENCNGSGSYVRKMIQDKEVAVKL-----LTKI--KEAEILSVLSHRNII 67  
 DB 190 DITMKHLGGGQYGEVYEVWKKYSLTVAVKTLKEDYVEEFLKEAAYMKIKHNLVQ 249  
 QY 68 FGVVILEPPNYGVIEYASLGSLVDYINSNRSEMDMHIMTWATDVAKGMVYLHEAV 127  
 DB 250 LAGVCTREPPFYITHEMYGNLIDLRKCNQEVAAVLLTMAQOISSAMEYLEKK-- 306  
 QY 128 KYIHRDLKSRNVIADGVLCIKDFGASRFNHTTMSLVGT-FP--WMAPEVIGSLPVS 184  
 DB 307 NFIIHRDLAARNCLGVSENNHVKVADFGLSLMTGDIYTAAGTKFPIKMTAPESLAVNKS 366  
 QY 185 ETCDDTYSYGVVLEMTREVP-FPKGLEQVAVLVEKNERLTIPSSCPSPFAELHQC 243  
 DB 367 IKSDVAFGVILMEIATYGMSPYIGDLSQV-YELLKDYRMERPEGCKEYVILMRACW 425  
 QY 244 EADAKKPSFKQI 256  
 DB 426 QWNPDSRPAFAEI 438

Sun May 4 10:01:18 2003

Search completed: May 1, 2003, 20:36:17  
Job time : 28.8452 secs

us-09-757-982-5.rsp

---

GenCore version 5.1.4 p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:28:47 ; Search time 50.7671 seconds

(without alignments)  
1846.697 Million cell updates/sec

Title: us-09-757-982-5

Perfect score: 2375

Sequence: 1 MSSIGASFVQIKFDDIQFFE.....GDDDDDDGGEDNDMDNSE 455

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

al number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_virus:\*

16: sp\_bacteriophage:\*

17: sp\_archaeop:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2375	100.0	455	4	Q9HCC4
2	2299.5	96.8	454	11	Q9ESL3
3	1751.5	73.7	800	4	Q9NYL2
4	1749.5	73.7	800	4	Q9NYS5
5	1748.5	73.6	800	4	Q9HDD2
6	1748.5	73.6	800	4	Q9NVE9
7	1743.5	73.4	802	11	Q9ESL4
8	1402.5	59.1	371	13	Q9QZV8
9	601	25.3	1066	4	Q9H2N5
10	598.5	25.2	1001	11	Q9VDG6
11	597.5	25.2	554	4	Q9H1Y7
12	590.5	24.9	1036	4	Q8WNN1
13	588.5	24.8	570	4	Q8WNN2
14	581.5	24.5	847	4	Q16584
15	580.5	24.4	850	11	Q9J315
16	572.5	24.1	740	5	Q21982

17	568.5	23.9	966	4	Q43283	043283 homo sapien
18	565.5	23.8	859	4	Q8WY25	Q8WY25 homo sapien
19	560.5	23.6	406	10	Q23719	Q23719 arabidopsis
20	560.5	23.6	880	10	Q8S9K4	Q8S9K4 arabidopsis
21	544.5	22.9	1020	5	Q9W313	Q9W313 drosophila
22	544.5	22.9	1148	5	Q9SVF6	Q9SVF6 drosophila
23	544.5	22.9	1161	5	Q95UN8	Q95UN8 drosophila
24	529.5	22.3	886	10	Q9LYI8	Q9LYI8 arabidopsis
25	527.5	22.2	977	5	Q9VW24	Q9VW24 drosophila
26	510.5	21.5	1020	10	Q9C9U5	Q9C9U5 arabidopsis
27	505	21.3	855	5	Q01700	Q01700 caenorhabdi
28	495	20.8	982	10	Q6S833	Q6S833 lycopersico
29	494	20.8	903	10	Q9FPR5	Q9FPR5 oryza sativ
30	483.5	20.4	847	10	Q93XL9	Q93XL9 roosa hybrid
31	483	20.3	491	4	Q9N270	Q9N270 homo sapien
32	480	20.2	829	10	Q24027	Q24027 lycopersico
33	478.5	20.1	933	10	Q9FPR3	Q9FPR3 arabidopsis
34	475.5	20.0	412	10	Q9M085	Q9M085 arabidopsis
35	475	20.0	957	10	Q9FPR4	Q9FPR4 hordeum vul
36	472.5	19.9	616	13	Q73613	Q73613 xenopus lae
37	471.5	19.9	606	11	Q923A8	Q923A8 mus musculu
38	471	19.8	777	10	Q9C833	Q9C833 arabidopsis
39	471	19.8	806	10	Q9ZSD8	Q9ZSD8 lycopersico
40	471	19.8	829	10	Q9ZSD9	Q9ZSD9 lycopersico
41	470.5	19.8	763	10	Q9C903	Q9C903 arabidopsis
42	470	19.8	735	10	Q93ZU3	Q93ZU3 arabidopsis
43	470	19.8	736	10	Q82754	Q82754 arabidopsis
44	468.5	19.7	678	5	Q9V306	Q9V306 drosophila
45	468	19.7	539	4	Q9NTRI	Q9NTRI homo sapien

## ALIGNMENTS

### RESULT 1

Q9HCC4 PRELIMINARY; PRT; 455 AA.

AC Q9HCC4; 01-MAR-2001 (T-EMBLrel. 16, Created)

DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)

DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)

DE MLTK-beta (Similar to sterile-alpha motif and leucine zipper containing kinase AZK) (Mixed lineage kinase) (Mixed lineage kinase-related kinase MRK-beta).

DE MLTK.

GN Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OC NCBI\_TaxID=9606;

OX [1]

RA SEQUENCE FROM N.A.

RA Gotoh T., Adachi M., Nishida E.;

RT "Identification and Characterization of a Novel MAP Kinase Kinase kinase, MLTK."

RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

RL [2]

RP SEQUENCE FROM N.A.

RP TISSUE-COLON;

RC Strausberg R.;

RA Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RL [3]

RP SEQUENCE FROM N.A.

RP Action S.;

RT "MLK-mixed lineage kinase."

RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RL [4]

RP SEQUENCE FROM N.A.

RP MEDLINE=21950776; PubMed=11836244;

RA Gross E.A., Callow M.G., Waldbaum L., Thomas S., Ruggieri R.;

RT "MRK, a Mixed Lineage Kinase-related Molecule That Plays a Role in Gamma-Radiation-Induced Cell Cycle Arrest."

RT J. Biol. Chem. 277:13873-13882(2002).

CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL; AB049734; BAB16445.1; -  
 DR EMBL; BC001401; AA01401.1; -  
 DR EMBL; AF325454; AAK11615.1; -  
 DR EMBL; AF480462; AAL85892.1; -  
 DR HSSP; P12931; 1FMK.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00221; STYKc; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; TYKc; 1.  
 DR PROSITE; PS50011; PROTEIN KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE\_ST; 1.  
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 455 AA; 51582 MW; B87DB844AD58B752 CRC64;

Query Match 100.0%; Score 2375; DB 4; Length 455;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-174; Mismatches 0; Indels 0; Gaps 0;  
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFPDLOFFNCGGSGSVYRAKWIQDKEVAVKLLKIEKEAIIISVL 60  
 DB 1 MSSLGASFVQIKFPDLOFFNCGGSGSVYRAKWIQDKEVAVKLLKIEKEAIIISVL 60  
 QY 61 SHRNIIQFYGVILLEPPNYGIVTEYASLGSLDYINSRSEEMDMHIMTWATDVAKGMHY 120  
 DB 61 SHRNIIQFYGVILLEPPNYGIVTEYASLGSLDYINSRSEEMDMHIMTWATDVAKGMHY 120  
 QY 121 LHMEAPVKVIHRDLKSNVVIADGVLIKDFGASRFHNTHTMSLVGTFFPMAPEVIQS 180  
 DB 121 LHMEAPVKVIHRDLKSNVVIADGVLIKDFGASRFHNTHTMSLVGTFFPMAPEVIQS 180  
 QY 181 LPVSTCDTYSYGVVLEWMLTREVPFKGLEGLQVAMLVVEKNERLTTPSSCPSPFAELIH 240  
 DB 181 LPVSTCDTYSYGVVLEWMLTREVPFKGLEGLQVAMLVVEKNERLTTPSSCPSPFAELIH 240  
 QY 241 QCEWADAKKRPSPFKQIISLESMSNDTSLPDKCNFLHNKAEWRCIEATLERLKLERD 300  
 DB 241 QCEWADAKKRPSPFKQIISLESMSNDTSLPDKCNFLHNKAEWRCIEATLERLKLERD 300  
 QY 301 LSPFKOEIKERERRLKWEOKLTQOSTNTPLLPLAARMSSESYFESKTEESNSAEMSCOI 360  
 DB 301 LSPFKOEIKERERRLKWEOKLTQOSTNTPLLPLAARMSSESYFESKTEESNSAEMSCOI 360  
 QY 361 TATNSGEGHGNPISLOAMMLMGFGDIFSMNKAGAVMHSQMGINMOAKONSSKTTSKRRGK 420  
 DB 361 TATNSGEGHGNPISLOAMMLMGFGDIFSMNKAGAVMHSQMGINMOAKONSSKTTSKRRGK 420  
 QY 421 KVNMLGFSDFDLSGDDDDDDGEEEDNDMNSE 455  
 DB 421 KVNMLGFSDFDLSGDDDDDDGEEEDNDMNSE 455

RESULT 2  
 Q9ESL3 PRELIMINARY; PRT; 454 AA.  
 ID Q9ESL3;  
 AC Q9ESL3;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE MLTK-Delta.  
 GN ZAK OR MLTK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=27264927; PubMed=11042189;

RA Gotoh I., Adachi M., Nishida E.;  
 RT "Identification and Characterization of a Novel Map Kinase Kinase  
 Kinase, MLTK.";  
 RL J. Biol. Chem., 276:4276-4286 (2001).  
 CC SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL; AB049732; BAB16443.1; -  
 DR HSSP; P12931; 1FMK.  
 DR MGD; MG11931274; Zak.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00221; STYKc; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR SMART; SM00219; TYKc; 1.  
 DR PROSITE; PS50011; PROTEIN KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE\_ST; 1.  
 DR ATP-binding; Serine/threonine-protein kinase; Transferase.  
 KW SEQUENCE 454 AA; 51366 MW; 35C2FCDD729D9395 CRC64;

Query Match 96.8%; Score 2299.5; DB 11; Length 454;  
 Best Local Similarity 96.5%; Pred. No. 7.1e-169; Mismatches 5; Indels 1; Gaps 1;  
 Matches 439; Conservative 10; Mismatches 5; Indels 1; Gaps 1;

QY 1 MSSLGASFVQIKFPDLOFFNCGGSGSVYRAKWIQDKEVAVKLLKIEKEAIIISVL 60  
 DB 1 MSSLGASFVQIKFPDLOFFNCGGSGSVYRAKWIQDKEVAVKLLKIEKEAIIISVL 60  
 QY 61 SHRNIIQFYGVILLEPPNYGIVTEYASLGSLDYINSRSEEMDMHIMTWATDVAKGMHY 120  
 DB 61 SHRNIIQFYGVILLEPPNYGIVTEYASLGSLDYINSRSEEMDMHIMTWATDVAKGMHY 120  
 QY 121 LHMEAPVKVIHRDLKSNVVIADGVLIKDFGASRFHNTHTMSLVGTFFPMAPEVIQS 180  
 DB 121 LHMEAPVKVIHRDLKSNVVIADGVLIKDFGASRFHNTHTMSLVGTFFPMAPEVIQS 180  
 QY 181 LPVSTCDTYSYGVVLEWMLTREVPFKGLEGLQVAMLVVEKNERLTTPSSCPSPFAELIH 240  
 DB 181 LPVSTCDTYSYGVVLEWMLTREVPFKGLEGLQVAMLVVEKNERLTTPSSCPSPFAELIH 240  
 QY 241 QCEWADAKKRPSPFKQIISLESMSNDTSLPDKCNFLHNKAEWRCIEATLERLKLERD 300  
 DB 241 QCEWADAKKRPSPFKQIISLESMSNDTSLPDKCNFLHNKAEWRCIEATLERLKLERD 300  
 QY 301 LSPFKOEIKERERRLKWEOKLTQOSTNTPLLPLAARMSSESYFESKTEESNSAEMSCOI 360  
 DB 301 LSPFKOEIKERERRLKWEOKLTQOSTNTPLLPLAARMSSESYFESKTEESNSAEMSCOI 360  
 QY 361 TATNSGEGHGNPISLOAMMLMGFGDIFSMNKAGAVMHSQMGINMOAKONSSKTTSKRRGK 420  
 DB 361 TATNSGEGHGNPISLOAMMLMGFGDIFSMNKAGAVMHSQMGINMOAKONSSKTTSKRRGK 420  
 QY 421 KVNMLGFSDFDLSGDDDDDDGEEEDNDMNSE 455  
 DB 421 KVNMLGFSDFDLSGDDDDDDGEEEDNDMNSE 455

RESULT 3  
 Q9NYL2 PRELIMINARY; PRT; 800 AA.  
 ID Q9NYL2;  
 AC Q9NYL2;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Mixed lineage kinase.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN (1)

RP SEQUENCE FROM N.A.  
 RX MEDLINE=20384179; PubMed=10924358;  
 RA Liu T.C., Huang C.J., Chu Y.C., Wei C.C., Chou C.C., Chou M.Y.,  
 RA Chou C.K., Yang J.J.;  
 RT "Cloning and expression of ZAK, a mixed lineage kinase-like protein  
 RT containing a leucine zipper and a sterile-alpha motif.";  
 RT Biochem. Biophys. Res. Commun. 274:811-816(2000).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AF238255; AAF63490.1; -  
 DR HSBP; P12931; IFMK.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001660; SAM.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00036; pkinase.1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase.1.  
 DR SMART; SM00454; SAM; 1.  
 DR SMART; SM00221; STYK; 1.  
 DR PROSITE; PSS0011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PSS00108; PROTEIN KINASE ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 800 AA; 91264 MW; DA82D7ABB2082F43 CRC64;

Query Match 73.7%; Score 1751.5; DB 4; Length 800;  
 Best Local Similarity 76.1%; Pred. No. 2.4e-126;  
 Matches 360; Conservative 24; Mismatches 56; Indels 33; Gaps 9;

QY 1 MSSLGASFYVQIKFDDLOFENCGGSGFSGSVYRAKMTISQDKEVAVVKLLKIEKEAELISVL 60  
 DB 1 MSSLGASFYVQIKFDDLOFENCGGSGFSGSVYRAKMTISQDKEVAVVKLLKIEKEAELISVL 60  
 QY 61 SHRNIIQFVGVILEPPNYGIVTEYASLSGLYDINSNREEMDMHMTATDVAKGMY 120  
 DB 61 SHRNIIQFVGVILEPPNYGIVTEYASLSGLYDINSNREEMDMHMTATDVAKGMY 120  
 QY 121 LHMEAPVKYIHRDLKSRNVIAADGVLCIDFGASRFNHTTHMSLVGTFPMMAPEVIQS 180  
 DB 121 LHMEAPVKYIHRDLKSRNVIAADGVLCIDFGASRFNHTTHMSLVGTFPMMAPEVIQS 180  
 QY 121 LHMEAPVKYIHRDLKSRNVIAADGVLCIDFGASRFNHTTHMSLVGTFPMMAPEVIQS 180  
 DB 121 LHMEAPVKYIHRDLKSRNVIAADGVLCIDFGASRFNHTTHMSLVGTFPMMAPEVIQS 180  
 QY 181 LPSVETCDTYSYGVVLEWMLTREVPFKGLEGLQVAMLVVEKNERLTISSCPBSFAELH 240  
 DB 181 LPSVETCDTYSYGVVLEWMLTREVPFKGLEGLQVAMLVVEKNERLTISSCPBSFAELH 240  
 QY 181 LPSVETCDTYSYGVVLEWMLTREVPFKGLEGLQVAMLVVEKNERLTISSCPBSFAELH 240  
 DB 181 LPSVETCDTYSYGVVLEWMLTREVPFKGLEGLQVAMLVVEKNERLTISSCPBSFAELH 240  
 QY 241 OCMEADAKKRPSPFOIISLESMSNDTSLPDKCNSFLHNKAEMWRCIEATLERLKLKLERD 300  
 DB 241 OCMEADAKKRPSPFOIISLESMSNDTSLPDKCNSFLHNKAEMWRCIEATLERLKLKLERD 300  
 QY 241 OCMEADAKKRPSPFOIISLESMSNDTSLPDKCNSFLHNKAEMWRCIEATLERLKLKLERD 300  
 DB 241 OCMEADAKKRPSPFOIISLESMSNDTSLPDKCNSFLHNKAEMWRCIEATLERLKLKLERD 300  
 QY 301 LSPKQELKERERRLKMEQKLTQESNTPLL-LPLAARMBEESYF--ESKTEESNSAE 355  
 DB 301 LSPKQELKERERRLKMEQKLTQESNTPLL-LPLAARMBEESYF--ESKTEESNSAE 355  
 QY 301 LSPKQELKERERRLKMEQKLTQESNTPLL-LPLAARMBEESYF--ESKTEESNSAE 355  
 DB 301 LSPKQELKERERRLKMEQKLTQESNTPLL-LPLAARMBEESYF--ESKTEESNSAE 355  
 QY 356 MSCQITATSNBEGHGMPSLOAMLMGFGDIFSMN--KAGAVNH--SGMQ-----INM-404  
 DB 356 MSCQITATSNBEGHGMPSLOAMLMGFGDIFSMN--KAGAVNH--SGMQ-----INM-404  
 QY 356 MSCQITATSNBEGHGMPSLOAMLMGFGDIFSMN--KAGAVNH--SGMQ-----INM-404  
 DB 356 MSCQITATSNBEGHGMPSLOAMLMGFGDIFSMN--KAGAVNH--SGMQ-----INM-404  
 QY 361 MSVYASLFKEN-----NITGRLLILLEEDLKMDGIVSKGHIHFPSAIEKLTHDYINLP 415  
 DB 361 MSVYASLFKEN-----NITGRLLILLEEDLKMDGIVSKGHIHFPSAIEKLTHDYINLP 415  
 QY 405 ---QAKONSSKTTSKRGKRVNMLGSPDPLSGDDDD-----DGEER 446  
 DB 405 ---QAKONSSKTTSKRGKRVNMLGSPDPLSGDDDD-----DGEER 446  
 QY 416 HFPPLIKDSGGEPEBNEKIVNLELVFG-FHLKPGTGPOCKKMYMEMDGE 467  
 DB 416 HFPPLIKDSGGEPEBNEKIVNLELVFG-FHLKPGTGPOCKKMYMEMDGE 467

RESULT 4  
 09HCC5 PRELIMINARY; PRT; 800 AA.  
 AC 09HCC5;  
 DT 01-MAR-2001 (TEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)  
 DE 01-MAR-2002 (TEMBLrel. 20, Last annotation update)  
 GN MTK-alpha.  
 OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21264927; PubMed=11042189;  
 RA Gotch I., Adachi M., Nishida E.;  
 RT "Identification and characterization of a novel MAP kinase kinase  
 RT kinase, MTK.";  
 RT J. Biol. Chem. 276:4276-4286(2001).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AB049733; BAB16444.1; -  
 DR HSBP; P12931; IFMK.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001660; SAM.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00036; pkinase.1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase.1.  
 DR SMART; SM00454; SAM; 1.  
 DR SMART; SM00221; STYK; 1.  
 DR SMART; SM00220; STYK; 1.  
 DR SMART; SM00219; TYRK; 1.  
 DR PROSITE; PSS0011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PSS00108; PROTEIN KINASE ST; 1.  
 KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 800 AA; 91188 MW; 2C8593824AB3FAD0 CRC64;

Query Match 73.7%; Score 1749.5; DB 4; Length 800;  
 Best Local Similarity 76.1%; Pred. No. 3.4e-126;  
 Matches 360; Conservative 23; Mismatches 57; Indels 33; Gaps 9;

QY 1 MSSLGASFYVQIKFDDLOFENCGGSGFSGSVYRAKMTISQDKEVAVVKLLKIEKEAELISVL 60  
 DB 1 MSSLGASFYVQIKFDDLOFENCGGSGFSGSVYRAKMTISQDKEVAVVKLLKIEKEAELISVL 60  
 QY 61 SHRNIIQFVGVILEPPNYGIVTEYASLSGLYDINSNREEMDMHMTATDVAKGMY 120  
 DB 61 SHRNIIQFVGVILEPPNYGIVTEYASLSGLYDINSNREEMDMHMTATDVAKGMY 120  
 QY 121 LHMEAPVKYIHRDLKSRNVIAADGVLCIDFGASRFNHTTHMSLVGTFPMMAPEVIQS 180  
 DB 121 LHMEAPVKYIHRDLKSRNVIAADGVLCIDFGASRFNHTTHMSLVGTFPMMAPEVIQS 180  
 QY 121 LHMEAPVKYIHRDLKSRNVIAADGVLCIDFGASRFNHTTHMSLVGTFPMMAPEVIQS 180  
 DB 121 LHMEAPVKYIHRDLKSRNVIAADGVLCIDFGASRFNHTTHMSLVGTFPMMAPEVIQS 180  
 QY 181 LPSVETCDTYSYGVVLEWMLTREVPFKGLEGLQVAMLVVEKNERLTISSCPBSFAELH 240  
 DB 181 LPSVETCDTYSYGVVLEWMLTREVPFKGLEGLQVAMLVVEKNERLTISSCPBSFAELH 240  
 QY 181 LPSVETCDTYSYGVVLEWMLTREVPFKGLEGLQVAMLVVEKNERLTISSCPBSFAELH 240  
 DB 181 LPSVETCDTYSYGVVLEWMLTREVPFKGLEGLQVAMLVVEKNERLTISSCPBSFAELH 240  
 QY 241 OCMEADAKKRPSPFOIISLESMSNDTSLPDKCNSFLHNKAEMWRCIEATLERLKLKLERD 300  
 DB 241 OCMEADAKKRPSPFOIISLESMSNDTSLPDKCNSFLHNKAEMWRCIEATLERLKLKLERD 300  
 QY 241 OCMEADAKKRPSPFOIISLESMSNDTSLPDKCNSFLHNKAEMWRCIEATLERLKLKLERD 300  
 DB 241 OCMEADAKKRPSPFOIISLESMSNDTSLPDKCNSFLHNKAEMWRCIEATLERLKLKLERD 300  
 QY 301 LSPKQELKERERRLKMEQKLTQESNTPLL-LPLAARMBEESYF--ESKTEESNSAE 355  
 DB 301 LSPKQELKERERRLKMEQKLTQESNTPLL-LPLAARMBEESYF--ESKTEESNSAE 355  
 QY 301 LSPKQELKERERRLKMEQKLTQESNTPLL-LPLAARMBEESYF--ESKTEESNSAE 355  
 DB 301 LSPKQELKERERRLKMEQKLTQESNTPLL-LPLAARMBEESYF--ESKTEESNSAE 355  
 QY 356 MSCQITATSNBEGHGMPSLOAMLMGFGDIFSMN--KAGAVNH--SGMQ-----INM-404  
 DB 356 MSCQITATSNBEGHGMPSLOAMLMGFGDIFSMN--KAGAVNH--SGMQ-----INM-404  
 QY 356 MSCQITATSNBEGHGMPSLOAMLMGFGDIFSMN--KAGAVNH--SGMQ-----INM-404  
 DB 356 MSCQITATSNBEGHGMPSLOAMLMGFGDIFSMN--KAGAVNH--SGMQ-----INM-404  
 QY 361 MSVYASLFKEN-----NITGRLLILLEEDLKMDGIVSKGHIHFPSAIEKLTHDYINLP 415  
 DB 361 MSVYASLFKEN-----NITGRLLILLEEDLKMDGIVSKGHIHFPSAIEKLTHDYINLP 415  
 QY 405 ---QAKONSSKTTSKRGKRVNMLGSPDPLSGDDDD-----DGEER 446  
 DB 405 ---QAKONSSKTTSKRGKRVNMLGSPDPLSGDDDD-----DGEER 446  
 QY 416 HFPPLIKDSGGEPEBNEKIVNLELVFG-FHLKPGTGPOCKKMYMEMDGE 467  
 DB 416 HFPPLIKDSGGEPEBNEKIVNLELVFG-FHLKPGTGPOCKKMYMEMDGE 467

RESULT 5  
 09HDD2 PRELIMINARY; PRT; 800 AA.  
 AC 09HDD2;  
 DT 01-MAR-2001 (TEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel). 16, last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel). 20, last annotation update)  
 DT Placible mixed-lineage kinase protein.  
 GN MUKLAK.  
 OS Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_Taxid=9606;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIMPHOID ORGAN;  
 RA Aze Y., Ueda N., lineage kinase derived from LAK cell."  
 RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AB030034; BAB12040.1; -  
 DR HSP; P12931; 1PMK.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001660; Ser\_thr\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00536; SAM; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR PRODOM; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00454; SAM; 1.  
 DR SMART; SM00221; STYK; 1.  
 DR SMART; SM00220; S\_TK; 1.  
 DR SMART; SM00219; TYRK; 1.  
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 KW SEQUENCE 800 AA; 9155 MW; B2814509EC54B07A CRC64;  
 SQ

Query Match 73.6%; Score 1748.5; DB 4; Length 800;  
 Best Local Similarity 76.1%; Pred. No. 4.1e-126;  
 Matches 360; Conservative 23; Mismatches 57; Indels 33; Gaps 9;

QY 1 MSLSGASFOVQIKFPDDLOFFENCAGSFGSYRAKMTISODKEVAVKLLKIEKEAETLSVL 60  
 DB 1 MSLSGASFOVQIKFPDDLOFFENCAGSFGSYRAKMTISODKEVAVKLLKIEKEAETLSVL 60  
 QY 61 SHRNIIQFYGVILPEPNYGIIVTEYASIGSLYDYNINRSSEMDHMTATVAKGMHY 120  
 DB 61 SHRNIIQFYGVILPEPNYGIIVTEYASIGSLYDYNINRSSEMDHMTATVAKGMHY 120  
 QY 121 LHMFAPVKVIHRDLKSRNVVIAADGVKICDPGASRFHNTHTMSLVGTFFPMAPEVIOS 180  
 DB 121 LHMFAPVKVIHRDLKSRNVVIAADGVKICDPGASRFHNTHTMSLVGTFFPMAPEVIOS 180  
 QY 121 LHMFAPVKVIHRDLKSRNVVIAADGVKICDPGASRFHNTHTMSLVGTFFPMAPEVIOS 180  
 DB 121 LHMFAPVKVIHRDLKSRNVVIAADGVKICDPGASRFHNTHTMSLVGTFFPMAPEVIOS 180  
 QY 181 LPVSETCDITYSGVVLWEMLTREVPFKGLBGLQVAMLVENKERTLTPSSCPSPFALLH 240  
 DB 181 LPVSETCDITYSGVVLWEMLTREVPFKGLBGLQVAMLVENKERTLTPSSCPSPFALLH 240  
 QY 241 QCEWADAKRRSPFOIITISLESMSNDTSLPDKNSFLHNKAEMCEIATELRLKLERD 300  
 DB 241 QCEWADAKRRSPFOIITISLESMSNDTSLPDKNSFLHNKAEMCEIATELRLKLERD 300  
 QY 301 LSFKEQELKERERRRLKMEQKLTQOSNTPLL-LPLAARMESESYF--ESKTEESNAE 355  
 DB 301 LSFKEQELKERERRRLKMEQKLTQOSNTPLL-LPLAARMESESYF--ESKTEESNAE 355  
 QY 301 LSFKEQELKERERRRLKMEQKLTQOSNTPLL-LPLAARMESESYF--ESKTEESNAE 355  
 DB 301 LSFKEQELKERERRRLKMEQKLTQOSNTPLL-LPLAARMESESYF--ESKTEESNAE 355  
 QY 356 MSCQITATNSGEGHGMNSLQAMLMGFGDI-FSNV--KAGAVM--SGMQ-----INM- 404  
 DB 356 MSCQITATNSGEGHGMNSLQAMLMGFGDI-FSNV--KAGAVM--SGMQ-----INM- 404  
 QY 361 NSVYASLPKEN-----NTGKRLILLEEDLDKNGIVSKGHIIFKSAIEKLTHTDYNLF 415  
 DB 361 NSVYASLPKEN-----NTGKRLILLEEDLDKNGIVSKGHIIFKSAIEKLTHTDYNLF 415  
 QY 405 ---QAKNSKSTTKRGRKGVAMALGFSDFLSSEGDDDD-----DGECE 446  
 DB 405 ---QAKNSKSTTKRGRKGVAMALGFSDFLSSEGDDDD-----DGECE 446  
 QY 416 HRPPLIKDGGEPRENEKIVLVLVFG-FHLKPGTGQDCMKMYMEMDDE 467  
 DB 416 HRPPLIKDGGEPRENEKIVLVLVFG-FHLKPGTGQDCMKMYMEMDDE 467

RESULT 6

Q9NYES PRELIMINARY; PRT; 800 AA.  
 ID Q9NYES  
 AC Q9NYES;  
 DT 01-OCT-2000 (TrEMBLrel). 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel). 15, last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel). 21, last annotation update)  
 DE Sterile-alpha motif and leucine zipper containing kinase AZK (Mixed  
 OS Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_Taxid=9606;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RA McNece J.U., Fima N., Diamond T.B., Dower S.K., Guesdon F.;  
 RT "Cloning and characterisation of AZK, a mixed lineage kinase  
 RT containing a sterile-alpha motif."  
 RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RL [2]  
 CC EMBL; AF251441; AAF65822.1; -  
 CC EMBL; AF480461; AAL65891.1; -  
 DR HSP; P12931; 1PMK.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001660; Ser\_thr\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00536; SAM; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR PRODOM; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00454; SAM; 1.  
 DR SMART; SM00221; STYK; 1.  
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 KW SEQUENCE 800 AA; 91181 MW; B289D836EC528295 CRC64;  
 SQ

Query Match 73.6%; Score 1748.5; DB 4; Length 800;  
 Best Local Similarity 76.1%; Pred. No. 4.1e-126;  
 Matches 360; Conservative 23; Mismatches 57; Indels 33; Gaps 9;

QY 1 MSLSGASFOVQIKFPDDLOFFENCAGSFGSYRAKMTISODKEVAVKLLKIEKEAETLSVL 60  
 DB 1 MSLSGASFOVQIKFPDDLOFFENCAGSFGSYRAKMTISODKEVAVKLLKIEKEAETLSVL 60  
 QY 61 SHRNIIQFYGVILPEPNYGIIVTEYASIGSLYDYNINRSSEMDHMTATVAKGMHY 120  
 DB 61 SHRNIIQFYGVILPEPNYGIIVTEYASIGSLYDYNINRSSEMDHMTATVAKGMHY 120  
 QY 121 LHMFAPVKVIHRDLKSRNVVIAADGVKICDPGASRFHNTHTMSLVGTFFPMAPEVIOS 180  
 DB 121 LHMFAPVKVIHRDLKSRNVVIAADGVKICDPGASRFHNTHTMSLVGTFFPMAPEVIOS 180  
 QY 121 LHMFAPVKVIHRDLKSRNVVIAADGVKICDPGASRFHNTHTMSLVGTFFPMAPEVIOS 180  
 DB 121 LHMFAPVKVIHRDLKSRNVVIAADGVKICDPGASRFHNTHTMSLVGTFFPMAPEVIOS 180  
 QY 181 LPVSETCDITYSGVVLWEMLTREVPFKGLBGLQVAMLVENKERTLTPSSCPSPFALLH 240  
 DB 181 LPVSETCDITYSGVVLWEMLTREVPFKGLBGLQVAMLVENKERTLTPSSCPSPFALLH 240  
 QY 241 QCEWADAKRRSPFOIITISLESMSNDTSLPDKNSFLHNKAEMCEIATELRLKLERD 300  
 DB 241 QCEWADAKRRSPFOIITISLESMSNDTSLPDKNSFLHNKAEMCEIATELRLKLERD 300  
 QY 301 LSFKEQELKERERRRLKMEQKLTQOSNTPLL-LPLAARMESESYF--ESKTEESNAE 355  
 DB 301 LSFKEQELKERERRRLKMEQKLTQOSNTPLL-LPLAARMESESYF--ESKTEESNAE 355  
 QY 301 LSFKEQELKERERRRLKMEQKLTQOSNTPLL-LPLAARMESESYF--ESKTEESNAE 355  
 DB 301 LSFKEQELKERERRRLKMEQKLTQOSNTPLL-LPLAARMESESYF--ESKTEESNAE 355



Qy 356 MSCQITATNSGEHGMNPSIQAMMLMGFGDI FSNM--KAGAVNH--SGNQ-----INM- 404  
 Db 361 MSYVASLFKEN-----NITGKRLILLEEDLKDMGIYSKGHIHFPSALEKHLVDINLF 415  
 Qy 405 ---QAKONSSTKTSKRGKRYNMALGSPDPLSGDDDD-----DGESE 446  
 Db 416 HFFPLIKDSGGEPEENEKEIYNLVLFG-FHLKPGTGPODCKMKRYMENDGDE 467

RESULT 7  
 Q9ESL4 PRELIMINARY; PRT; 802 AA.

ID Q9ESL4  
 AC Q9ESL4;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE MLTK alpha.  
 GN ZAK OR MLTK.  
 OS Mus musculus (Mouse).  
 NC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21264927; PubMed=11042189;  
 RA Gotch J., Adachi M., Nishida E.;  
 RT Identification and Characterization of a Novel MAP Kinase Kinase  
 RL Kinase, MLTK.;  
 CC J. Biol. Chem. 276:4276-4286(2001).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AB049731; BAB16442.1; -  
 DR HSP; P12931; IPMK.  
 DR MGD; MGI:1931274; Zak.  
 DR InterPro; IPR000194; ATPase\_a/bcentre.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001660; SAM.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00536; SAM; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00454; SAM; 1.  
 DR SMART; SM00221; STYK; 1.  
 DR SMART; SM00220; S\_TKC; 1.  
 DR SMART; SM00219; TYKC; 1.  
 DR PROSITE; PS00152; ATPASE\_ALPHA\_BETA; UNKNOWN\_1.  
 DR PROSITE; PS00101; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 802 AA; 91719 MW; DA3IDPF8312M43CC CRC64;

Query Match 73.4%; Score 1743.5; DB 11; Length 802;  
 Best Local Similarity 74.3%; Pred. No. 1e-125;  
 Matches 365; Conservative 20; Mismatches 58; Indels 45; Gaps 7;

Qy 1 MSSLGASFVQIKFPDLOPFENCGGSGFSGVYRAKVISODKEVAVKLIKIEKEAIIISVL 60  
 Db 1 MSSLGASFVQIKFPDLOPFENCGGSGFSGVYRAKVISODKEVAVKLIKIEKEAIIISVL 60  
 Qy 61 SHRNIIQFYGVILEPPNGIVTEYASLSGLYDINSNRSEEMDMHIMTATVAKGMHY 120  
 Db 61 SHRNIIQFYGVILEPPNGIVTEYASLSGLYDINSNRSEEMDMHIMTATVAKGMHY 120  
 Qy 121 LHMEAPVKVIHRDLKSRNVVIAADGVLIKIDFGASRFHNHTTHMSLVGTFFPMAPEVIQS 180  
 Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLIKIDFGASRFHNHTTHMSLVGTFFPMAPEVIQS 180  
 Qy 181 LPVSETCDTYSYGVVLMEMLTREVFPKGLGLOVAMLVVEKNERLITPSSCPSPFAELIH 240  
 Db 181 LPVSETCDTYSYGVVLMEMLTREVFPKGLGLOVAMLVVEKNERLITPSSCPSPFAELIH 240

Qy 241 QCEWADAKRPSFKQIISILSMNSDITSLPDKNSFLHNKAEMCEIATLERLKLKLERD 300  
 Db 241 QCEWADAKRPSFKQIISILSMNSDITSLPDKNSFLHNKAEMCEIATLERLKLKLERD 300  
 Qy 301 LSPKEQELKERERRLKWMEOULTQOSNTPIL--LPLAARSESEYF-----SEKTE 349  
 Db 301 LSPKEQELKERERRLKWMEOULTQOSNTPILPSPFEIQAWEEDVYFWVQOLVRKSESSVE 360  
 Qy 350 ESNSEMSQITATNSGEHGMNPSIQAMMLMGFGDI FSNM--KAGAVNHSGNQINNQAQ 407  
 Db 361 MSYVASLFKENITG-----KRLILLEEDLKDMGIYSKGHIHFPSALEKHLVDINLF 409  
 Qy 408 -----QNSKTSKRGKRYNMALGSPDPLSGDDDD-----DGESE 446  
 Db 410 DYNLHFFPLIKDSGGEPEENEKEIYNLVLFG-FHLKPGTGPODCKMKRYMENDGDE 467

RESULT 8  
 Q90ZY8 PRELIMINARY; PRT; 371 AA.

ID Q90ZY8  
 AC Q90ZY8;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Protein kinase Npk.  
 OS Brachydanio rerio (Zebrafish) (Zebra danio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 NC NCBI\_TaxId=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Chou C.-M., Lee I.-L., Leu J.-H., Huang C.-J.;  
 RT "A novel protein kinase, znpk, from the zebrafish."  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AF265343; AAK52416.1; -  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
 KW ATP-binding; Kinase; Transferase.  
 SQ SEQUENCE 371 AA; 42456 MW; 9B918B8AB200296 CRC64;

Query Match 59.1%; Score 1402.5; DB 13; Length 371;  
 Best Local Similarity 77.6%; Pred. No. 6.4e-100;  
 Matches 266; Conservative 32; Mismatches 38; Indels 7; Gaps 2;

Qy 1 MSSLGASFVQIKFPDLOPFENCGGSGFSGVYRAKVISODKEVAVKLIKIEKEAIIISVL 60  
 Db 32 MSSLGASFVQIKFPDRIFFENCGGSGFSGVYRAHVPODKEVAVKLIKIDKEAIIISVL 91  
 Qy 61 SHRNIIQFYGVILEPPNGIVTEYASLSGLYDINSNRSEEMDMHIMTATVAKGMHY 120  
 Db 92 SHRNIIQFYGVILEPPNGIVTEYASLSGLYDINSNRSEEMDMHIMTATVAKGMHY 151  
 Qy 121 LHMEAPVKVIHRDLKSRNVVIAADGVLIKIDFGASRFHNHTTHMSLVGTFFPMAPEVIQS 180  
 Db 152 LHMEAPVKVIHRDLKSRNVVIAADGVLIKIDFGASRFHNHTTHMSLVGTFFPMAPEVIQS 211  
 Qy 181 LPVSETCDTYSYGVVLMEMLTREVFPKGLGLOVAMLVVEKNERLITPSSCPSPFAELIH 240  
 Db 212 LPVSETCDTYSYGVVLMEMLTREVFPKGLGLOVAMLVVEKNERLITPSSCPSPFAELIH 271  
 Qy 241 QCEWADAKRPSFKQIISILSMNSDITSLPDKNSFLHNKAEMCEIATLERLKLKLERD 300  
 Db 272 QCEWADAKRPSFKQIISILSMNSDITSLPDKNSFLHNKAEMCEIATLERLKLKLERD 331  
 Qy 301 LSPKEQELKERERRLKWMEOULTQOSN--TPLLPLAARSESE 341  
 Db 332 LSPKEQELKERERRLKWMEOULTQOSN--TPLLPLAARSESE 369

## RESULT 9

Q942N5 PRELIMINARY; PRT; 1066 AA.  
 ID Q942N5  
 AC Q942N5  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Mixed lineage kinase MLK1 (Fragment).  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NCBI\_TaxID=9606;  
 [1]

SEQUENCE FROM N.A.  
 RA McEneaney J.T., Dower S.K., Guesdon F.;  
 RT "cDNA sequence and gene organisation of mixed lineage kinase 1."  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 DR EMBL; AF251442; AAC44591.1; -.  
 DR HSP; P29355; ISEM.

DR InterPro; IPR000719; Ser\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_kinase; 1.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR SMART; SM00220; S\_TKC; 1.  
 DR SMART; SM00219; Tyrc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR KINASE; SH3 domain.  
 KW NON TER  
 FT  
 SQ SEQUENCE 1066 AA; 118463 MW; EDD08EBE7482723 CRC64;

Query Match 25.3%; Score 601; DB 4; Length 1066;

Best Local Similarity 39.2%; Pred. No. 1.4e-37;  
 Matches 135; Conservative 60; Mismatches 117; Indels 32; Gaps 5;

QY 9 VOIKFDLQFFENGCGGSGFSGVYRAKMIISODKEAVK-----KLKIEKAEI 56  
 85 LEIDFELTLEELTIGGFGKVRATWID--EVAVKARHPDEDISQTIENVRQEAEL 142  
 DB 57 LSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDINSNSSEMDMHIMTAVAK 116  
 143 FALVKKHPNIIALRGVCLKEPNLCLVMEFARGPLNRLVLSGRK--IPPILVMAVQAKR 199  
 QY 117 GMYLMEAPVAVKVIHRDLKSRNVI-----AADGLKICDFGASRPHNHTHSLVG 168  
 DB 200 GMYLMEAPVAVKVIHRDLKSRNVI-----AADGLKICDFGASRPHNHTHSLVG 259  
 QY 169 TFPMAPEVIOQLPVSFETCDTYSYGVVLMEMLTREVPEFKGLQVAMLVKNERLTP 228  
 DB 260 TYAMMAPEVIRASWFSQVWSYGVVLMELTGEVPEFRGIDGLAVAYGVAMNKALPLP 319  
 QY 229 SSCPRFAELLHQCEWADAKKRPFKOIIISLSMSNDTSLPKKCSFLNKAKEMRCETE 288  
 DB 320 STCEPEAPKLMEDCWNDPSPRPFTILDOULTTEESGFEMPKOSFGLQDNKHEIO 379  
 QY 289 ATTERLKLKEDLSFKEOEL-----KERERLKMWEOKLTREQ 325  
 DB 380 RMFDQRAKEKELTWEELTRALQKNGEBELLRRREGELAR 423

RESULT 10  
 Q8VDG6

ID Q8VDG6 PRELIMINARY; PRT; 1001 AA.

AC Q8VDG6  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Similar to mitogen-activated protein kinase kinase 9.  
 OS Mus musculus (mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 [1]

SEQUENCE FROM N.A.  
 RA Struhsberg R.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC021891; AAH21891.1; -.  
 DR InterPro; IPR000719; Euk\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR ProDom; PD000001; Euk\_kinase; 1.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR SMART; SM00220; S\_TKC; 1.  
 DR SMART; SM00219; Tyrc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
 DR KINASE.  
 KW KINASE.  
 SQ SEQUENCE 1001 AA; 109983 MW; E10042C868B9953C CRC64;

Query Match 25.2%; Score 598.5; DB 11; Length 1001;

Best Local Similarity 33.7%; Pred. No. 2e-37;  
 Matches 161; Conservative 85; Mismatches 161; Indels 71; Gaps 16;

QY 9 VOIKFDLQFFENGCGGSGFSGVYRAKMIISODKEAVK-----KLKIEKAEI 56  
 DB 103 VHDVFERLELKEILIGAGFGQVYRATW--QGCEVAVKARHPEDDAAAASVREAREL 160  
 QY 57 LSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDINSNSSEMDMHIMTAVAK 106  
 DB 161 FALVKKHPNIIALRGVCLKEPNLCLVMEFARGPLNRLVLSGRK--IPPILVMAVQAKR 220  
 QY 107 IMTAVTAVAGMYLHMEAPVAVKVIHRDLKSRNVI-----AADGLKICDFGASRPH 158  
 DB 221 LVNMAVOIARGLYLHEEAVVPIHRDLKSNLLEKIEHDDICNKTILKITDGLAREW 280  
 QY 159 NHTHMSLVGTFFPMAPEVIOQLPVSFETCDTYSYGVVLMEMLTREVPEFKGLQVAMLV 218  
 DB 281 HRTTMSAGTYAMMAPEVIRASWFSQVWSYGVVLMELTGEVPEFRGIDGLAVAYGV 340  
 QY 219 VEKNERLTIPSSCPRAELLHQCEWADAKKRPFKOIIISLSMSNDTSLPKKCSFLNK 276  
 DB 341 AVNKLTLPIPTCEPEAPKLMEDCWNDPSPRPFTILDOULTTEESGFEMPKOSFGLQ 398  
 QY 277 LNKAKEMRCETEATLERLKLKEDLSFKEOEL-----KERERLKMWEOKLTREQ 329  
 DB 399 HSMQDWTLEIQMSF--LRTKEKELRSREELSPALQKSGQELRLRRRCQOLAEREIDV 457  
 QY 330 L--LLPLAAMSEBS--YFESKTEESNSAEMSCQITATNSGEHGNN--PS-----LOA 377  
 DB 458 LERELNVLIFOLSGEAPVAVKVRKGRFRGRRLK-----DGRIRISLPSPDFOKITVOA 510  
 QY 378 MMLMGFGDIFSNKAKGAVWHSQ--MQLNMAKONSSTKTSKRGRKVNMAAGSDFD 432  
 DB 511 SPTL--DKRRSSDSGLCSPPSGSLMLPRLAQILDTSPENNKTKGR--NVVFRQEDPE 563

RESULT 11  
 Q9H1Y7

ID 09HLY PRELIMINARY; PRT; 564 AA.  
 AC 09HLY;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE D362P8.3 (Similar to MAPK10 (Mitogen-activated protein kinase kinase 10)) (Fragment).  
 DE D362P8.3.  
 GN D362P8.3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OC NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hall R.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 DR EMBL; AL13380; CAC17571.1; -.  
 DR HSPSP; P06241; ISHF.  
 DR InterPro; IPR000719; Euk\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00069; kinase; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_kinase; 1.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SMO0220; SH3; 1.  
 DR SMART; SMO0220; S\_TKc; 1.  
 DR SMART; SMO0219; TyKc; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00101; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS00002; SH3; 1.  
 KW Kinase; SH3 domain.  
 KW NON\_TER  
 SQ SEQUENCE 564 AA; 62400 MW; D4P91E14842E7EBB CRC64;  
 Query Match 25.2%; Score 597.5; DB 4; Length 564;  
 Best Local Similarity 37.4%; Pred. No. 1.1e-37;  
 Matches 135; Conservative 66; Mismatches 115; Indels 45; Gaps 7;  
 QY 6 ASFVQIKFDDIQFENCSSGSGSVYRAKWIISQKEVAVKLLK-----IEKE 53  
 DB 114 SSPVHAFFERLEIKELIGAGFGQVYRATW--QGQEVAVKAAQDPQDAAAASVRR 171  
 QY 54 AELISVLSHRNIQFYGVILEPPNYGIVTEVASLSGLYDYI-----NSNRSEE 101  
 DB 172 ARLFAMLRHFNITELRGVCIQOPHLCLVLEFARGALNRLAANAAAPDPRAPRRARR 231  
 QY 102 MMDHIMTATDVAKGMHYLHMEAPVKVTHRDLSKNVY-----AADGVLCIDFG 153  
 DB 232 IPRHVLVNMVAQIARGLYLHEAFVPIIHRDLKSSNILLLEIEHDDICNKTIKTIDFG 291  
 QY 154 ASRPHNHTTMSLVGTFPMAPEVIQSLPVSETCDTVSYGVVLMEMLTREVFPKGLGLQ 213  
 DB 292 LAEMHRTTKMSTAGTAYMAWAPVYIKSLFSKGSIDWSGVLMELLTGEVPYRGIDGLA 351  
 QY 214 VAVLVYKNERLTPSSCRSPFAELHQCWEADAKRPSFKQIISLESMSND--TSLPD 271  
 DB 352 VAVVAVNKLTLPIPTCEPFAELKMECWQODPHIRPSFALILEQTLTAEGAVMTMPQ 411  
 QY 272 KCSFLNKAEMRCCEIATLERLKKLERDLSFKEQEL-----KERERLKMWEQKLT 324  
 DB 412 E--SFHSMQDWMKLEIQMFPDELRTKELRSREBELTRALAQKSEBELKREQLAE 469  
 QY 325 Q 325  
 DB 470 R 470

RESULT 12  
 08WMN1 PRELIMINARY; PRT; 1036 AA.  
 ID 08WMN1;  
 AC 08WMN1;  
 DT 01-MAR-2002 (Tremblrel. 20, Created)  
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Mixed lineage kinase 4beta.  
 GN MK4BETA.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OC NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Krishna S., Protopopov A., Rynditch A., Zabarovsky E., Kashuba V.;  
 RT "MK4, a new member of mixed lineage kinases."  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ111798; CAC84640.1; -.  
 DR InterPro; IPR000719; Euk\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00069; kinase; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR ProDom; PD000001; Euk\_kinase; 1.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SMO0326; SH3; 1.  
 DR SMART; SMO0220; S\_TKc; 1.  
 DR SMART; SMO0219; TyKc; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN\_1.  
 DR PROSITE; PS00101; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
 DR PROSITE; PS00002; SH3; 1.  
 KW Kinase.  
 KW NON\_TER  
 SQ SEQUENCE 1036 AA; 113809 MW; B9C2ACF397C2CF37 CRC64;  
 Query Match 24.9%; Score 590.5; DB 4; Length 1036;  
 Best Local Similarity 33.1%; Pred. No. 8.7e-37;  
 Matches 149; Conservative 75; Mismatches 147; Indels 79; Gaps 9;  
 QY 6 ASFVQIKFDDIQFENCSSGSGSVYRAKWIISQKEVAVKLLK-----IEKE 53  
 DB 114 SSPVHAFFERLEIKELIGAGFGQVYRATW--QGQEVAVKAAQDPQDAAAASVRR 171  
 QY 54 AELISVLSHRNIQFYGVILEPPNYGIVTEVASLSGLYDYI-----NSNRSEE 101  
 DB 172 ARLFAMLRHFNITELRGVCIQOPHLCLVLEFARGALNRLAANAAAPDPRAPRRARR 231  
 QY 102 MMDHIMTATDVAKGMHYLHMEAPVKVTHRDLSKNVY-----AADGVLCIDFG 153  
 DB 232 IPRHVLVNMVAQIARGLYLHEAFVPIIHRDLKSSNILLLEIEHDDICNKTIKTIDFG 291  
 QY 154 ASRPHNHTTMSLVGTFPMAPEVIQSLPVSETCDTVSYGVVLMEMLTREVFPKGLGLQ 213  
 DB 292 LAEMHRTTKMSTAGTAYMAWAPVYIKSLFSKGSIDWSGVLMELLTGEVPYRGIDGLA 351  
 QY 214 VAVLVYKNERLTPSSCRSPFAELHQCWEADAKRPSFKQIISLESMSND--TSLPD 271  
 DB 352 VAVVAVNKLTLPIPTCEPFAELKMECWQODPHIRPSFALILEQTLTAEGAVMTMPQ 411  
 QY 272 KCSFLNKAEMRCCEIATLERLKKLERDLSFKEQEL-----KERERLKMWEQKLT 324  
 DB 412 E--SFHSMQDWMKLEIQMFPDELRTKELRSREBELTRALAQKSEBELKREQLAE 469  
 QY 325 Q 325  
 DB 470 REIDVLERELNITLIPOLNQEKPVKKRGKFKRSRLKLDGHRISLPDFQKITVQASP 529  
 QY 355 EM-----SCQITATSNQEGGNPISLOAMML 380  
 DB 530 NUDKRRSLNSSSSPPSSPTMPRLRLAQL 559



QY 121 LHMAFPVVIHRDLKSNVV---IAADV---LKICDFGASRFHNTHTMSLVGTFPM 172  
 DB 229 LHCEALVPVIRHDLKSNVILLOPIBGDDMEHKTUKITDFGLAREMHTTQMSAAGTYAM 288  
 QY 173 MAPEVIOSLPSECTDYSGVVLWEMLTREVPFKLEGLOVAVLVKNERLTIPSSCP 232  
 DB 289 MAPEVIKASTSKSDVWSFGVLLWELLTGEVPRGIDCLAVAGVAVNKLTLPIPTCP 348  
 QY 233 RSFAELHQCWEADAKKPSFKQIISIESMSNDTSLPDKNSFLHNKAEMRCIEATLE 292  
 DB 349 EPFAQLMADCMADPHRRPDPFASILQOLEALAEQVLEMPRDSFHSQEGWKREIQLGFD 408  
 QY 293 RLKLERDLSFKOE-----LKERERLKMWEOKLTQOSNTPLL 331  
 DB 409 ELRAKEKELLSREBELTRAAEGRSQAEQLRRREHLLAQWELVEFVERDVTLLL 461

## RESULT 15

09J15

PRELIMINARY; PRT; 850 AA.

NC 09J15; 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Mixed lineage kinase 3.  
 GN MLK3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurigmachi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/OLA;  
 RX MEDLINE=20354997; PubMed=10894943;  
 RA Saridaki A., Ferraz C., Demaille J., Scherer G., Roux A.-F.;  
 RT "Genomic sequencing reveals the structure of the Kcnk6 and Mapk11  
 RT genes and their close vicinity to the Sipa1 gene on mouse chromosome  
 RT 19."  
 RL Cytogenet. Cell Genet. 89:85-88(2000).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 DR EMBL; AF151142; AAF73281.1; -.  
 DR HSSP; P29355; 1SEM.  
 DR InterPro; IPR000719; Euk\_Pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_Pkinase.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001245; Tyr\_Pkinase.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR PRINTS; PRO0018; SH3; 1.  
 DR PRINTS; PRO0452; SH3DOMAIN.  
 DR PRINTS; PRO0109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_Pkinase; 1.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR SMART; SM00219; TyrcK; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS0011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR ATP-binding; Kinase; SH3 domain; Serine/threonine-protein kinase;  
 KW Transferrase.  
 SQ SEQUENCE 850 AA; 9319 MW; 8F026CB352DC10E CRC64;

## Query Match

Best Local Similarity 24.4%; Score 580.5; DB 11; Length 850;  
 Matches 102; Conservative 59; Mismatches 123; Indels 39; Gaps 7;

QY 13 FDDIQFENCGSGSGSVYRAKWSIQ-----DKFAVYKLLKIEKAEIISLVLS 61  
 DB 115 FOELRLBEVIGIGFGKVGSRGELVAVYARQDPDEDISV-TASVQGEARLPAMLA 173  
 QY 62 HRNIIQFYGVILRPNNGIYEVASLGSLYDINSNRSEMDMHIH-TWATDVAKGMHY 120

DB 174 HPIITIALAVCIIEENLCLVMEIYAAAGGPLSPALAGRVP-----PHVLVNWAVQIARGMHY 229  
 QY 121 LHMAFPVVIHRDLKSNVV---IAADV---LKICDFGASRFHNTHTMSLVGTFPM 172  
 DB 230 LHCEALVPVIRHDLKSNVILLOPIBGDDMEHKTUKITDFGLAREMHTTQMSAAGTYAM 289  
 QY 173 MAPEVIOSLPSECTDYSGVVLWEMLTREVPFKLEGLOVAVLVKNERLTIPSSCP 232  
 DB 290 MAPEVIKASTSKSDVWSFGVLLWELLTGEVPRGIDCLAVAGVAVNKLTLPIPTCP 349  
 QY 233 RSFAELHQCWEADAKKPSFKQIISIESMSNDTSLPDKNSFLHNKAEMRCIEATLE 292  
 DB 350 EPFAQLMADCMADPHRRPDPFASILQOLEALAEQVLEMPRDSFHSQEGWKREIQLGFD 409  
 QY 293 RLKLERDLSFKOE-----LKERERLKMWEOKLTQOSNTPLL 331  
 DB 410 ELRAKEKELLSREBELTRAAEGRSQAEQLRRREHLLAQWELVEFVERDVTLLL 462

Search completed: May 1, 2003, 20:37:45  
 Job time : 58.7671 secs





FT	Modified-site	/note= "protein kinase ATP-binding site"
FT	/note= 159	
FT	Modified-site	/note= "potential glycosylation site"
FT	/note= 234	
FT	Modified-site	/note= "potential phosphorylation site"
FT	/note= 252	
FT	Modified-site	/note= "potential phosphorylation site"
FT	/note= 258	
FT	Modified-site	/note= "potential phosphorylation site"
FT	/note= 265	
FT	Modified-site	/note= "potential glycosylation site"
FT	/note= 268	
FT	Region	/note= "potential phosphorylation site"
FT	/note= 294...322	
FT	/note= 302	"leucine zipper"
FT	Modified-site	/note= "potential phosphorylation site"
FT	/note= 302	
FT	Modified-site	/note= "potential phosphorylation site"
FT	/note= 342	
FT	Modified-site	/note= "potential phosphorylation site"
FT	/note= 343	
FT	Modified-site	/note= "potential phosphorylation site"
FT	/note= 346	
FT	Modified-site	/note= "potential phosphorylation site"
FT	/note= 364	
FT	Modified-site	/note= "potential phosphorylation site"
FT	/note= 409	
FT	Modified-site	/note= "potential glycosylation site"
FT	/note= 410	
FT	Modified-site	/note= "potential phosphorylation site"
FT	/note= 414	
FT	Modified-site	/note= "potential phosphorylation site"
FT	/note= 415	
FT	Modified-site	/note= "potential phosphorylation site"
FT	/note= 429	
FT	Modified-site	/note= "potential phosphorylation site"
FT	/note= 434	
FT	Modified-site	/note= "potential phosphorylation site"
PN	WO20005532-A2.	
PD	21-SEP-2000.	
XX	17-MAR-2000; 2000WO-US07277.	
PF	18-MAR-1999; 99US-0125593.	
XX	20-MAY-1999; 99US-0135049.	
PR	09-JUL-1999; 99US-0143188.	
PA	(INCY-) INCTE PHARM INC.	
PI	Bandman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y,	
PI	Lu DAM, Au-Yang J;	
XX	WPI: 2000-602121/57.	
DR	N-PDSB; AAA5674.	
XX	Novel human intracellular phosphorylation regulator polypeptides and	
FT	polynucleotides for diagnosis, prevention and treatment of	
FT	neurological, cell proliferative and autoimmune/inflammatory disorders	
PT	-	
XX	Claim 1; Page 75-76; 96pp; English.	
PS		
CC	The present sequence represents a human regulator of intracellular	
CC	phosphorylation (HRIP). HRIP is useful for screening agonists and	
CC	antagonists of HRIP polypeptide. HRIP and its agonist or antagonist	
CC	are useful for treating a disease or condition associated with	
CC	decreased or increased expression of functional HRIP. Diseases treated	
CC	or diagnosed include neurological disorders such as stroke, Parkinson's	
CC	disease, demyelinating diseases, bacterial and viral meningitis and	
CC	other developmental disorders of the central nervous system,	

	neurovascular disorders, myasthenia gravis, cell proliferative disorders
CC	such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer
CC	including leukemia, melanoma, myeloma and cancer of the adrenal gland,
CC	bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/
CC	inflammatory disorder such as Addison's disease, acquired
CC	immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,
CC	rheumatoid arthritis, microbial infection and trauma.
XX	
SQ	Sequence      455 AA;
	Query Match                  100.0%; Score 2375; DB 21; Length 455;
	Best Local Similarity        100.0%; Pred. No. 6.5e-207;
	Matches    455; Conservative    0; Mismatches    0; Indels     0; Gaps     0;
QY	1 MSSIGASFVQIKFDDLOPFENCGGGSPGSVTRAKVISODKXAVAYKKLIKIEKEAIIISVL 60
Db	1 MSSIGASFVQIKFDDLOPFENCGGGSPGSVTRAKVISODKXAVAYKKLIKIEKEAIIISVL 60
QY	61 SHRNIIOFYGVILPEPNYGIVTEIASLSGLDYDINSNSEEMDMHIMTMAIDVAKGMHY 120
Db	61 SHRNIIOFYGVILPEPNYGIVTEIASLSGLDYDINSNSEEMDMHIMTMAIDVAKGMHY 120
QY	121 LHMEAPVAVIHRDLKSRRNVIAADGVLKICDPGARFRFNHTTHMSLVGFPMMADEVIOS 180
Db	121 LHMEAPVAVIHRDLKSRRNVIAADGVLKICDPGARFRFNHTTHMSLVGFPMMADEVIOS 180
QY	181 LPVSECTDTYSYGVVLMMLTREVPFKGLEQLQVAMLYVEKNERTITSSCPRSFAELIH 240
Db	181 LPVSECTDTYSYGVVLMMLTREVPFKGLEQLQVAMLYVEKNERTITIPSSCPRSFAELIH 240
QY	241 QCWEDAKKRPSFOITISILSMSNDTSLPKCNSEFLINKAEWRCEIATELRLKKLEKD 300
Db	241 QCWEDAKKRPSFOITISILSMSNDTSLPDKCNSFLINKAEWRCEIATELRLKKLERD 300
QY	301 LSFKEQELKERERRLKWMEOKLTBQSNTPLLLPLAARMSBSSEFSKTEESNSAEMSCOI 360
Db	301 LSFKEQELKERERRRKWMEOKLTBQSNTPLLLPLAARMSBSSEFSKTEESNSAEMSCOI 360
QY	361 TATSNGBEGHGMPISIQAMLNGFGDI FSNMRYAGAVMHSGMOINMAKONSSKTTSKRGRK 420
Db	361 TATSNGBEGHGMPISIQAMLNMGFDIFSNMRYAGAVMHSGMOINMAKONSSKTTSKRGRK 420
QY	421 KVNMMALGPSDFLSEGDDDDDDGEEENDMDNSE 455
Db	421 KVNMMALGPSDFLSEGDDDDDDGEEENDMDNSE 455
<hr/>	
RESULT 2	
AAy83278	
ID    AAy83278 standard; Protein: 455 AA.	
XX    AAy83278;	
AC    AAy83278;	
XX    DT    16-AUG-2000 (first entry)	
DE    Human survival regulating kinase (SRK).	
XX    Survival regulating kinase; SRK; Raf; MAPK; BAD; MAPKK; MAPKK;	
KW    HAX-1; protein kinase; autophosphorylation; cell growth; regulation;	
KW    apoptosis; cell survival; nuclear targeting; tumour; human;	
KW    autoimmune disease.	
XX    Homo sapiens.	
OS    Homo sapiens.	
XX    MO200022142-A2.	
PD    20-APR-2000.	
PF    20-SEP-1999;      99WO-US22008.	
PR    13-OCT-1998;      98US-0104088.	
PA    (ONTX-) ONTX PHARM INC.	



XX Ruggieri R, Collow M, Diaz P;  
XX WPI; 2000-317994/27.  
DR N-PSDB; AA293783.  
XX Novel human survival regulating kinase polypeptide for screening agents  
PT which modulate biological pathways associated with SRK useful in  
XX treating autoimmune diseases, tumors and apoptosis-related disorders  
PS Claim 4; Figure 2; 62pp; English.

CC Survival regulating kinases (SRK) are a class of proteins involved in  
CC cell signal transduction pathways such as mitogen-activated protein  
CC kinase pathways. A protein kinase activity means that the SRK can  
CC catalyze a reaction in which a phosphate group is transferred from a  
CC phosphate donor to a phosphate acceptor amino acid residue,  
CC preferably the hydroxyl side chain of a serine or threonine.  
CC Substrates for SRK include SRK, MAP and BAD and SRK's protein kinase  
CC activity is similar to that of a MAPKK such as Raf. has a range of  
CC other activities including a cell growth-regulatory activity, a cell  
CC survival promoting activity, a HAX-1 binding activity, an apoptosis  
CC suppressing activity a MAPKK activation or stimulatory activity, a  
CC nuclear targeting activity and a SRK-specific immunogenic activity.  
CC SRK is useful for identifying agents which modulate cellular  
CC transformations mediated by Ras and SRK and agents that modulate the  
CC apoptosis suppression activity of SRK. This information may be useful  
CC in the treatment of autoimmune diseases, tumors and apoptosis  
XX related disorders.

XX Sequence 455 AA;

Query Match 100.0%; Score 2375; DB 21; Length 455;  
Best Local Similarity 100.0%; Pred. No. 6.5e-207;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFWQIKFDDLOFFENCGGSGSVYRAKWIISODKEVAAYVKLIKIKKEAIIISVL 60  
DB 1 MSSLGASFWQIKFDDLOFFENCGGSGSVYRAKWIISODKEVAAYVKLIKIKKEAIIISVL 60  
QY 61 SHRNIIOFYGVILEPPNIGIYEVASLSGLYDINSNRSEEMDMHMTATDVAKGMHY 120  
DB 61 SHRNIIOFYGVILEPPNIGIYEVASLSGLYDINSNRSEEMDMHMTATDVAKGMHY 120  
QY 121 LHMEAPVAVIHRDLKSRNVVIAADGVLIKIDFGASRFNHTTHSLVGTFPWMAPEVIQS 180  
DB 121 LHMEAPVAVIHRDLKSRNVVIAADGVLIKIDFGASRFNHTTHSLVGTFPWMAPEVIQS 180  
QY 181 LPVSETCTYSYGVLWEMLTREVPFKGLEGLQVAMLVENKNERLTJPSSCPSPSFAELH 240  
DB 181 LPVSETCTYSYGVLWEMLTREVPFKGLEGLQVAMLVENKNERLTJPSSCPSPSFAELH 240  
QY 241 QCMWADAKKRSPFQOIIISILESMSNDTSLPDKCNSFLHNKAEMRCETATLERLKKLERD 300  
DB 241 QCMWADAKKRSPFQOIIISILESMSNDTSLPDKCNSFLHNKAEMRCETATLERLKKLERD 300  
QY 301 LSPFEOELKEKERERLTKMBOEKLTEQSNTPILLPLAARMSSESYFESTTESNSAEMSCOI 360  
DB 301 LSPFEOELKEKERERLTKMBOEKLTEQSNTPILLPLAARMSSESYFESTTESNSAEMSCOI 360  
QY 361 TATNSGEGHGMNPSLQAMMLMGFGDIFSNMKAAGVVMHSGMGMQIMQAKONSKTTSKRGRK 420  
DB 361 TATNSGEGHGMNPSLQAMMLMGFGDIFSNMKAAGVVMHSGMGMQIMQAKONSKTTSKRGRK 420  
QY 421 KVNMAAGFSPDFLSEGGDDDDDDGEEEDNDMDNSE 455  
DB 421 KVNMAAGFSPDFLSEGGDDDDDDGEEEDNDMDNSE 455

RESULT 3  
AA84321  
ID AAY84321 standard; protein; 455 AA.  
XX

AC AAY84321;  
XX 12-JUN-2000 (first entry)  
DT  
XX  
XX  
DE A human cardiovascular system associated protein kinase-2.  
XX  
KW Human; cardiovascular system associated protein kinase-2; CSAPK-2;  
KW signalling pathway; cell growth; cell differentiation; gene mapping;  
KW tissue typing; forensic identification; cardiovascular disease;  
KW congestive heart failure; transgenic animal.

OS Homo sapiens.

PN W0200014212-A1.

PD 16-MAR-2000.

PF 09-SEP-1999; 99MO-US20631.

PR 09-SEP-1998; 98US-0099657.

PR 23-SEP-1998; 98US-0163115.

PA (MILL-) MILLENNIUM PHARM INC.

PI Acton S;

DR WPI; 2000-271053/23.

DR N-PSDB; AA29726; AA29727.

PT New nucleic acid encoding cardiovascular system associated protein  
PT kinase, used e.g. for diagnosis, treatment and prevention of  
PT cardiovascular disease

PS Claim 2; Fig 2; 163pp; English.

CC The present sequence represents a human cardiovascular system associated  
CC protein kinase-2 (CSAPK-2). CSAPK polypeptides are involved in signalling  
CC pathways associated with cell growth and differentiation. The CSAPK  
CC polypeptides and polynucleotides are used to screen for agents that  
CC specifically modulate CSAPK, which are potential therapeutic agents.  
CC They are also used for diagnosis, prognosis or monitoring of  
CC CSAPK-related diseases, gene mapping, tissue typing and forensic  
CC identification, and for treating or preventing disorders associated  
CC with aberrant CSAPK expression or activity, especially cardiovascular  
CC diseases such as congestive heart failure. They can also be used in  
CC pharmacogenomics. The CSAPK polynucleotide may also be used to generate  
CC transgenic animals.

XX Sequence 455 AA;

Query Match 100.0%; Score 2375; DB 21; Length 455;  
Best Local Similarity 100.0%; Pred. No. 6.5e-207;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFWQIKFDDLOFFENCGGSGSVYRAKWIISODKEVAAYVKLIKIKKEAIIISVL 60  
DB 1 MSSLGASFWQIKFDDLOFFENCGGSGSVYRAKWIISODKEVAAYVKLIKIKKEAIIISVL 60  
QY 61 SHRNIIOFYGVILEPPNIGIYEVASLSGLYDINSNRSEEMDMHMTATDVAKGMHY 120  
DB 61 SHRNIIOFYGVILEPPNIGIYEVASLSGLYDINSNRSEEMDMHMTATDVAKGMHY 120  
QY 121 LHMEAPVAVIHRDLKSRNVVIAADGVLIKIDFGASRFNHTTHSLVGTFPWMAPEVIQS 180  
DB 121 LHMEAPVAVIHRDLKSRNVVIAADGVLIKIDFGASRFNHTTHSLVGTFPWMAPEVIQS 180  
QY 181 LPVSETCTYSYGVLWEMLTREVPFKGLEGLQVAMLVENKNERLTJPSSCPSPSFAELH 240  
DB 181 LPVSETCTYSYGVLWEMLTREVPFKGLEGLQVAMLVENKNERLTJPSSCPSPSFAELH 240  
QY 241 QCMWADAKKRSPFQOIIISILESMSNDTSLPDKCNSFLHNKAEMRCETATLERLKKLERD 300  
DB 241 QCMWADAKKRSPFQOIIISILESMSNDTSLPDKCNSFLHNKAEMRCETATLERLKKLERD 300

QY 301 LSPKEQLKREERRLKWEQKLTQOSTPILPLPLAARMSESYFESKTESNSAEMSCOI 360  
 DB 301 LSPKEQLKREERRLKWEQKLTQOSTPILPLPLAARMSESYFESKTESNSAEMSCOI 360  
 QY 361 TATNSGEGHGNPDLQAMLMGFGDIFSNKAGAVHSGMQLMOKAKONSKTTSKRGG 420  
 DB 361 TATNSGEGHGNPDLQAMLMGFGDIFSNKAGAVHSGMQLMOKAKONSKTTSKRGG 420  
 QY 421 KVNNALGFSDFDLSGGDDDDDDGEEEDNDMDNSE 455  
 DB 421 KVNNALGFSDFDLSGGDDDDDDGEEEDNDMDNSE 455  
 RESULT 4  
 AAM25322  
 ID AAM25322 standard; Protein; 473 AA.  
 AC AAM25322;  
 16-OCT-2001 (first entry)  
 DE Human protein sequence SEQ ID NO:837.  
 KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
 KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
 KW antibacterial; endocrine; cardiac; central nervous system; vitruclide;  
 KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;  
 KW antiagregant; haemostatic; vulnerary; antilucer; osteopathic; eczema;  
 KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;  
 KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
 KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;  
 KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
 KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
 KW genetic disease; haematopoietic disorder; platelet disorder; asthma;  
 KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;  
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;  
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
 KW neurological disorder.  
 OS Homo sapiens.  
 PN WO200153455-A2.  
 PD 26-JUL-2001.  
 PF 22-DEC-2000; 2000MO-US35017.  
 KY 23-DEC-1999; 99US-0471275.  
 PR 21-JAN-2000; 2000US-0486725.  
 PA 25-APR-2000; 2000US-0552317.  
 PI (HYSE-) \*HYSEQ INC.  
 PI Tang YT, Liu C, Drmanac RT,  
 DR WPI, 2001-457603/49.  
 DR N-PSDB; AAH99263.  
 XX  
 XX  
 PT Isolated human polynucleotides encoding polypeptides, useful for the  
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -  
 XX  
 XX  
 XX Claim 20, Page 191, 1217pp; English.  
 PS  
 PS AAH99166 to AAH99904 encode the human proteins given in AAM25225 to  
 CC AAM25963. The proteins can have activities based on the tissues and  
 CC cells they are expressed in, such as: antiinflammatory; antirheumatic;  
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;  
 CC central nervous system; vitruclide; anti-HIV; fungicide; antilucer;  
 CC cardiovascular; antianaemic; antiagregant; haemostatic; vulnerary;  
 CC antilucer; osteopathic; dermatological; antiallergic; antiasthmatic;  
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;  
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides

CC encoding them can be used in gene therapy, antisense therapy and vaccine  
 CC production. The proteins and polynucleotides are useful for screening for  
 CC agonists or antagonists of a protein and for the treatment and diagnosis  
 CC of disorders associated with the activity of a protein e.g. inflammation,  
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,  
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
 CC neurological disorders.  
 SQ Sequence 473 AA;  
 Query Match 99.6%; Score 2366; DB 22; Length 473;  
 Best Local Similarity 99.8%; Pred. No. 4,66-206;  
 Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MSISGASFYQIFKFDLPENCGGSGFSGVYRAKWIISQDEKVAVKKLIKREAEILSVL 60  
 DB 19 MSISGASFYQIFKFDLPENCGGSGFSGVYRAKWIISQDEKVAVKKLIKREAEILSVL 78  
 QY 61 SHRNIIQFYGVILPEPNYGIETEVASLSGLVDYINSNSEEKMDHITWATDVAKGMHY 120  
 DB 79 SHRNIIQFYGVILPEPNYGIETEVASLSGLVDYINSNSEEKMDHITWATDVAKGMHY 138  
 QY 121 LHMBAVKTHIDIKRNVAADGVKICDPGASRFNHTHMSLVGTFPMMADEVIOS 180  
 DB 139 LHMBAVKTHIDIKRNVAADGVKICDPGASRFNHTHMSLVGTFPMMADEVIOS 198  
 QY 181 LPVSETCDTYSYGVVLMEMLTREVPKGLGLQVAMLVKERNKRLTTPSCCRSPAEILH 240  
 DB 199 LPVSETCDTYSYGVVLMEMLTREVPKGLGLQVAMLVKERNKRLTTPSCCRSPAEILH 258  
 QY 241 QCWEADAKRPSFKQIISLESMSNTSLPDKNSFTLHKARMCIEATLRLKLEED 300  
 DB 259 QCWEADAKRPSFKQIISLESMSNTSLPDKNSFTLHKARMCIEATLRLKLEED 318  
 QY 301 LSPKEQLKREERRLKWEQKLTQOSTPILPLPLAARMSESYFESKTESNSAEMSCOI 360  
 DB 319 LSPKEQLKREERRLKWEQKLTQOSTPILPLPLAARMSESYFESKTESNSAEMSCOI 378  
 QY 361 TATNSGEGHGNPDLQAMLMGFGDIFSNKAGAVHSGMQLMOKAKONSKTTSKRGG 420  
 DB 379 TATNSGEGHGNPDLQAMLMGFGDIFSNKAGAVHSGMQLMOKAKONSKTTSKRGG 438  
 QY 421 KVNNALGFSDFDLSGGDDDDDDGEEEDNDMDNSE 455  
 DB 439 KVNNALGFSDFDLSGGDDDDDDGEEEDNDMDNSE 473  
 RESULT 5  
 AAB71957  
 ID AAB71957 standard; Protein; 800 AA.  
 AC AAB71957;  
 DT 11-MAY-2001 (first entry)  
 DE Human TGF-beta receptor encoded by cDNA clone HDP5M48.  
 KW Human; antileukemic; dermatological; immunosuppressive; cytostatic;  
 KW antiinflammatory; anti-HIV; immunostimulant; cardiac; vascular;  
 KW opthalmological; neuroprotective; nootropic; anticonvulsant; vaccine;  
 KW antiparkinsonian; antidiabetic; vulnerary; gene therapy; infection;  
 KW transforming growth factor; TGF; TGF-beta receptor; immune disorder;  
 KW hyperproliferative disorder; cardiovascular disease; angiogenesis;  
 KW neurological disorder.  
 OS Homo sapiens.  
 PN WO200112670-A1.

XX 22-FEB-2001.  
 XX 10-AUG-2000; 2000MO-US21736.  
 XX 13-AUG-1999; 99US-0148682.  
 PR 20-SEP-1999; 99US-0154887.  
 XX (HUMA-) HUMAN GENOME SCT INC.  
 XX Ruben SM, NI J;  
 PI WPI; 2001-202858/20.  
 DR N-PSDB; AAF75336.  
 XX Nucleic acid molecules encoding 12 transforming growth factor-beta  
 PT receptor polypeptides, useful for preventing, diagnosing and treating  
 PT e.g. cancers, Parkinson's disease and diabetic retinopathy -  
 XX

Claim 11; Page 293-295; 311pp; English.

The present sequence is one of 12 novel human transforming growth factor (TGF)-beta receptor polypeptides. The TGF-beta receptor polynucleotides and polypeptides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate polypeptide expression. Such diseases include immune disorders (e.g. multiple sclerosis, systemic lupus erythematosus and human immunodeficiency virus (HIV) infections), hyperproliferative disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases (e.g. Schmittar syndrome, Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic disorders (e.g. corneal graft neovascularization and diabetic retinopathy), neurological disorders (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease) and infectious diseases. The polynucleotides and polypeptides are also useful for promoting wound healing, regeneration and/or chemotaxis. The polynucleotides and their complementary sequences may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples. The polypeptides may be used as antigens in the production of antibodies and in assays to identify modulators of protein expression and activity. The anti-TGF-beta receptor antibodies may be used to down regulate expression and activity and as diagnostic agents for detecting the presence of the polypeptides in samples.

Sequence 800 AA;

Query Match 73.6%; Score 1748.5; DB 22; Length 800;  
 Best Local Similarity 76.1%; Pred. No. 1.3e-149;  
 Matches 360; Conservative 23; Mismatches 57; Indels 33; Gaps 9;

1 MSSLSGASVVOIKPDLOFPENCGGSGSVYRAKWTISODKEVAVKLIKIKERAEIISVL 60  
 DB 1 MSSLSGASVVOIKPDLOFPENCGGSGSVYRAKWTISODKEVAVKLIKIKERAEIISVL 60  
 QY 61 SHRNIQFYGVILPEPPNGIVTEVASLGLVDYINSRSEEMDMHMTATDVAKGMHY 120  
 DB 61 SHRNIQFYGVILPEPPNGIVTEVASLGLVDYINSRSEEMDMHMTATDVAKGMHY 120  
 QY 121 LHMEAPVAVIHRDLKSRNVIAADGVLIKIDFGASRFNHTTHSLVGTFFPMAPEVIQS 180  
 DB 121 LHMEAPVAVIHRDLKSRNVIAADGVLIKIDFGASRFNHTTHSLVGTFFPMAPEVIQS 180  
 QY 181 LPVSETCOTYGVAVLWMLTRVPPKGLBELOVAVLVEKNERLITPSSCPSPFAELIH 240  
 DB 181 LPVSETCOTYGVAVLWMLTRVPPKGLBELOVAVLVEKNERLITPSSCPSPFAELIH 240  
 QY 241 QCMWADAKKRPSPFOIISILESMNDTSLPDKNSPLHNKAMPCCEIATELERKJLERD 300  
 DB 241 QCMWADAKKRPSPFOIISILESMNDTSLPDKNSPLHNKAMPCCEIATELERKJLERD 300  
 QY 301 LSFKEQELKEKERLKWMEQKLTQOSNTPLL-LPLAARMSSESYF--ESKTEESNSAE 355  
 DB 301 LSFKEQELKEKERLKWMEQKLTQOSNTPLL-LPLAARMSSESYF--ESKTEESNSAE 355

QY 356 MSCOTTATSNGBGHQMPISLOAMLMGFGDIFSMN--KAGAVMH--SGMQ-----INM- 404  
 DB 361 MSVYASLFEKEN-----NITGRILLLEBEDJXMDGIYSKHIIIFKAIKELCTHDYINLP 415  
 QY 405 ---QAKNSKTTSKRKGKYNMALGFSDDPDLISGDDDD-----DQSEE 446  
 DB 416 HPPPLIKDSGGEPEBNEKEIVLELVFG-FHLKGTGTGPQDCKWIKWYMEMDDGE 467

# RESULT 6

AAB65673  
 ID AAB65673 standard; Protein; 800 AA.

AC AAB65673;

DT 27-MAR-2001 (first entry)

DE Novel protein kinase, SEQ ID NO: 201.

XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;  
 KW immunosuppressive; cardiant; renal; antiinflammatory; antisthmatic;  
 KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;  
 KW immune disorder; cardiovascular disease; neurodegenerative disease;  
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;  
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis.

OS Homo sapiens.

PN W0200073469-A2.

PD 07-DEC-2000.

PF 26-MAY-2000; 2000MO-US14842.

PR 28-MAY-1999; 99US-0136503.

XX (SUGEN-) SUGEN INC.

PI Plowman GD, Martinez R, Whyte D, Sudersanam S;

DR WPI; 2001-032161/04.

DR N-PSDB; AAF44701.

PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
 PT treating immune-related diseases and disorders, cardiovascular disease,  
 PT neurodegenerative diseases and/or cancers.

PS Claim 10; Fig 1; 310pp; English.

XX The present sequence is a novel protein kinase. The novel protein kinases  
 CC and the nucleic acids that encode them may be used in the treatment and  
 CC diagnosis of diseases associated with inappropriate kinase expression  
 CC such as immune-related diseases and disorders, cardiovascular disease,  
 CC neurodegenerative diseases and/or cancers. The nucleic acids and  
 CC complementary sequences may also be used as DNA probes in diagnostic  
 CC assays. The kinase polypeptides may be used as antigens in the production  
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies  
 CC and kinase antagonists may also be used to down regulate kinase  
 CC expression and activity. Diseases related to kinase expression and  
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune  
 CC disorders, complications of organ transplantation, myocardial infarction,  
 CC immune disorders, cardiomyopathies, strokes, renal failure,  
 CC oxidative-stress related disorders, chronic inflammatory bowel disease,  
 CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,  
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and  
 CC reproductive disorders.

Sequence 800 AA;

Query Match 73.6%; Score 1748.5; DB 22; Length 800;  
 Best Local Similarity 76.1%; Pred. No. 1.3e-149;  
 Matches 360; Conservative 23; Mismatches 57; Indels 33; Gaps 9;

QY 1 MSSIGASFWQIKFPDLOFFENCGGSGFSGSVYRAKMTISODKEVAVKTLKIEKAEITLSVL 60  
 1 MSSIGASFWQIKFPDLOFFENCGGSGFSGSVYRAKMTISODKEVAVKTLKIEKAEITLSVL 60  
 QY 61 SHRNIIOFYGVILLPPNYGIVTEYASLGSLYDYINSRSEMDMDHMTATDVAKGMHY 120  
 61 SHRNIIOFYGVILLPPNYGIVTEYASLGSLYDYINSRSEMDMDHMTATDVAKGMHY 120  
 QY 121 LHMEAPVKVTHRDLSKSNVVIADGVLTICDFGASRFPHNTTHMSLVGTPEPMAPEVIQS 180  
 121 LHMEAPVKVTHRDLSKSNVVIADGVLTICDFGASRFPHNTTHMSLVGTPEPMAPEVIQS 180  
 QY 181 LPVSETCDTYSYGVLWEMLTREVPFKGLEQVAMLVYKNERLTTPSSCPSPFAELH 240  
 181 LPVSETCDTYSYGVLWEMLTREVPFKGLEQVAMLVYKNERLTTPSSCPSPFAELH 240  
 QY 241 QCWEADAKKRPSPFKQIISILSMSNDTSLPDKNSFLHNKAEMRCETATLERLKLKLERD 300  
 241 QCWEADAKKRPSPFKQIISILSMSNDTSLPDKNSFLHNKAEMRCETATLERLKLKLERD 300  
 Db 301 LSPFKEOELKEBERRLKMWEOKLTESNTPL--LPLAARMESESYF--ESKTEESNSAE 355  
 301 LSPFKEOELKEBERRLKMWEOKLTESNTPL--LPLAARMESESYF--ESKTEESNSAE 355  
 QY 356 MSCQITATSNGBGGMPSLOAMLMGFQDIFSMN--KAGAVM--SGMQ-----IM- 404  
 356 MSCQITATSNGBGGMPSLOAMLMGFQDIFSMN--KAGAVM--SGMQ-----IM- 404  
 Db 361 MSVASFLEN-----NITGKRLILLEBEDLMDGIVSKGHIHFKSAIEKLTHTDYNLF 415  
 361 MSVASFLEN-----NITGKRLILLEBEDLMDGIVSKGHIHFKSAIEKLTHTDYNLF 415  
 QY 405 ---QAKONSKTTSKRGKQVMALGFSDPDLSEGDGDD-----DDGSE 446  
 405 ---QAKONSKTTSKRGKQVMALGFSDPDLSEGDGDD-----DDGSE 446  
 Db 416 HPPPLIKDSGGEPENBEKIYVLELVFG-FHLKPGTGQDCKMKMYMEMDGE 467  
 416 HPPPLIKDSGGEPENBEKIYVLELVFG-FHLKPGTGQDCKMKMYMEMDGE 467

RESULT 7  
 AAG75571  
 ID AAG75571 standard; Protein; 349 AA.

AC AAG75571;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen protein SEQ ID NO:6335.

XX Human, colon cancer; colon cancer antigen; diagnosis; detection;

KM colorectal carcinoma.

OS Homo sapiens.

XX WO200122920-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000WO-US26524.

XX 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCT INC.

PA Ruben SM, Barash SC, Bires CE, Rosen CA;

PI WPI, 2001-235357/24.

DR N-PSDB; AAH34976.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,

PT useful for preventing, diagnosing and/or treating colorectal cancers -

XX Claim 11; Page 7789-7790; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon

CC cancer-associated nucleic acid molecules (N) and proteins (P), where

CC the proteins are collectively known as colon cancer antigens. The colon

CC cancer antigens have cytostatic activity and can be used in gene

CC therapy and vaccine production. N and P may be used in the prevention,  
 CC diagnosis and treatment of diseases associated with inappropriate P  
 CC expression. For example, N and P may be used to treat disorders  
 CC associated with decreased expression by rectifying mutations or deletions  
 CC in a patient's genome that affect the activity of P by expressing P.  
 CC Inactive proteins or to supplement the patients own production of P.  
 CC Additionally, N may be used to produce the colon cancer-associated Ps,  
 CC by inserting the nucleic acids into a host cell and culturing the cell  
 CC to express the proteins. N and P can be used in the prevention, diagnosis  
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
 CC present invention.  
 CC and AAH37789 represent sequences used in the exemplification of the  
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
 CC missing at time of publication, meaning no sequences are present for  
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 349 AA;

Query Match 59.3%; Score 1409; DB 22; Length 349;

Best Local Similarity 100.0%; Pred. No. 2.5e-119;

Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSIGASFWQIKFPDLOFFENCGGSGFSGSVYRAKMTISODKEVAVKTLKIEKAEITLSVL 60  
 57 MSSIGASFWQIKFPDLOFFENCGGSGFSGSVYRAKMTISODKEVAVKTLKIEKAEITLSVL 116

Db 61 SHRNIIOFYGVILLPPNYGIVTEYASLGSLYDYINSRSEMDMDHMTATDVAKGMHY 120  
 117 SHRNIIOFYGVILLPPNYGIVTEYASLGSLYDYINSRSEMDMDHMTATDVAKGMHY 176

QY 121 LHMEAPVKVTHRDLSKSNVVIADGVLTICDFGASRFPHNTTHMSLVGTPEPMAPEVIQS 180  
 177 LHMEAPVKVTHRDLSKSNVVIADGVLTICDFGASRFPHNTTHMSLVGTPEPMAPEVIQS 236

QY 181 LPVSETCDTYSYGVLWEMLTREVPFKGLEQVAMLVYKNERLTTPSSCPSPFAELH 240  
 237 LPVSETCDTYSYGVLWEMLTREVPFKGLEQVAMLVYKNERLTTPSSCPSPFAELH 296

Db 241 QCWEADAKKRPSPFKQIISILSMSNDTSL 269  
 297 QCWEADAKKRPSPFKQIISILSMSNDTSL 325

RESULT 8  
 AAG03583  
 ID AAG03583 standard; Protein; 141 AA.

AC AAG03583;

DT 06-OCT-2000 (first entry)

DE Human secreted protein, SEQ ID NO: 7664.

XX Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KM gene therapy; chromosome mapping.

OS Homo sapiens.

XX EPI033401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST ) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI, 2000-500381/45.

XX N-PSDB; AAC03589.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 XX  
 PS Claim 13; SEQ ID 7664; 71bp + CD-ROM; English.

CC The present sequence is a polypeptide encoded by one of a large number  
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
 CC different tissues. EST sequences usually correspond mainly to the 3'  
 CC untranslated region (UTR) of the mRNA because they are often obtained  
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 CC those cases where longer cDNA sequences have been obtained, the full 5'  
 CC ends are rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 CC ends and can therefore be used to obtain full length cDNAs and genomic  
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 CC chromosome mapping procedures. They are used to obtain upstream  
 CC regulatory sequences and to design expression and secretion vectors.

Sequence 141 AA;

Query Match 31.3%; Score 744; DB 21; Length 141;  
 Best Local Similarity 98.6%; Pred. No. 1.6e-59;  
 Matches 139; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 102 MDMDHMTATDVAKGMYHMEAPVKVIRHDKSRNVYIADGVAKICDPGASRPHNNT 161  
 DB 1 MDMDHMTATDVAKGMYHMEAPVKVIRHDKSRNVYIADGVAKICDPGASRPHNNT 60  
 QY 162 THNSLVGTFPMMAPEVIOSLPVSETCDTYSYGVLWEMLTREVPFGLEGLOVAMLVVEK 221  
 DB 61 THNSLVGTFPMMAPEVIOSLPVSETCDTYSYGVLWEMLTREVPFGLEGLOVAMLVVEK 120  
 QY 222 NERLITPSSCGRSPFALLHQC 242  
 DB 121 NERLITPSSCGRSPFALLHQC 141

RESULT 9

AAB65552  
 ID AAB65552 standard; Protein; 124 AA.

AC AAB65552;

DT 27-MAR-2001 (first entry)

C-terminal specific to novel human protein kinase MLK4B.

Human; protein kinase; antiarthritic; antisclerotic; immunosuppressive;  
 KW cardiant; renal; antiinflammatory; antidiabetic; osteopathic;  
 KW dermatological; antidiabetic; antileukemia; gene therapy; vaccine;  
 KW immune disorder; cardiovascular disease; neurodegenerative disease;  
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;  
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis.

Homo sapiens.

WO200073469-A2.

07-DEC-2000.

26-MAY-2000; 2000WO-US14842.

28-MAY-1999; 99US-0136503.

(SUGC-) SUGEN INC.

PI Plowman GD, Martinez R, Whyte D, Sudersanam S;

DR WPI; 2001-032161/04.

PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and

PT treating immune-related diseases and disorders, cardiovascular disease,  
 PT neurodegenerative diseases and/or cancers -

XX Example 1; Page 116; 310pp; English.

CC The present sequence is given in a specification relating to novel  
 CC protein kinases. The protein kinases and the nucleic acids that encode  
 CC them may be used in the treatment and diagnosis of diseases  
 CC associated with inappropriate kinase expression such as immune-related  
 CC diseases and disorders, cardiovascular disease, neurodegenerative  
 CC diseases and/or cancers. The nucleic acids and complementary sequences  
 CC may also be used as DNA probes in diagnostic assays. The kinase  
 CC polypeptides may be used as antigens in the production of antibodies of  
 CC kinase expression and activity. Anti-kinase antibodies and kinase  
 CC antagonists may also be used to down regulate kinase expression and  
 CC activity. Diseases related to kinase expression and activity include  
 CC rheumatoid arthritis, atherosclerosis, autoimmune disorders,  
 CC complications of organ transplantation, myocardial infarction, immune  
 CC disorders, cardiomyopathies, strokes, renal failure, oxidative-stress  
 CC related disorders, chronic inflammatory bowel disease, chronic  
 CC inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis,  
 CC psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive  
 CC disorders.

Sequence 124 AA;

Query Match 26.9%; Score 638; DB 22; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-50;  
 Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 332 LPLAARSESYPESTEESNSAEMSCOTTATSNAGCHGNBLSQMLMGFGDIFSNMK 391  
 DB 1 LPLAARSESYPESTEESNSAEMSCOTTATSNAGCHGNBLSQMLMGFGDIFSNMK 60  
 QY 392 AGAVHSGMQINNQAOKNSKTTSKRGKKNVMAIGFSPDISEGDDDDDDGEEEDNDM 451  
 DB 61 AGAVHSGMQINNQAOKNSKTTSKRGKKNVMAIGFSPDISEGDDDDDDGEEEDNDM 120  
 QY 452 DNSE 455  
 DB 121 DNSE 124

RESULT 10

ABP61000  
 ID ABP61000 standard; Protein; 1021 AA.

AC ABP61000;

DT 10-SEP-2002 (first entry)

DE Novel human protein. SEQ ID 87.

Human; cytosolic; vulnery; antiarteriosclerotic; antiparkinsonian;  
 KW neurotrophic; neuroprotective; immunosuppressive; haemostatic;  
 KW antiinflammatory; cardiant; antitumor; vitinide; antithyroid;  
 KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;  
 KW wound healing disorders; atherosclerosis; Parkinson's disease;  
 KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;  
 KW inflammation; neoplastic disease; nervous system disorder;  
 KW cardiovascular disorders; pancreatitis; respiratory disorder;  
 KW hyperproliferation; systemic autoimmune disease; hyper-immunity;  
 KW developmental abnormality; gastrointestinal ulceration; neuropathy;  
 KW haematological disease; metabolic disease; sperm dysfunction;  
 KW thyroid disorder; hypothyroidism; brain damage; colitis;  
 KW cone photo- transduction deficiency; neurological disease; stroke;  
 KW angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;  
 KW trachea; lymph node; muscular system; obesity; anorexia;  
 KW growth abnormality; precocious puberty.

Homo sapiens.

WO200250105-A1.



CC lateral sclerosis, viral infections, infections caused by prions,  
 CC bacteria and fungi, ocular diseases, migraines, pain, sexual dysfunction,  
 CC mood disorders, attention disorders, cognition disorders, hypotension,  
 CC hypertension, psychotic disorders, neurological disorders, dyskinesias,  
 CC metabolic disorders, and organ transplant rejection. They are also useful  
 CC for treating rhinitis, autoimmune diseases, osteoarthritis, psoriasis,  
 CC osteoarthritis, asthma, chronic inflammatory pelvic disease, chronic  
 CC inflammatory bowel disease, rheumatoid arthritis, metabolic disorders  
 CC such as diabetes, obesity, cardiovascular diseases such as reperfusion  
 CC injury, coronary thrombosis, clotting disorders and atherosclerosis.  
 CC ocular diseases such as glaucoma, retinopathy and macular degeneration,  
 CC psychiatric and neurological disorders such as anxiety, schizophrenia,  
 CC dementia, manic depression, etc. The polynucleotides are useful in gene  
 CC therapy techniques to treat the above mentioned disorders. Sequences  
 CC AAB85491-85522 represent the human protein kinases of the invention.

Sequence 719 AA;

Query Match 25.2%; Score 599.5; DB 22; Length 719;  
 Best Local Similarity 33.3%; Pred. No. 2,6e-45;

Matches 150; Conservative 75; Mismatches 146; Indels 79; Gaps 9;

QY 6 ASFVQIFEDDLOFFENCGGSGFSVYRAKWIISQKEVAAYKLIK-----IEKE 53  
 DB 114 SSVVHVAFERLELKELGAGFGQVYRATW--QGGEVAVKAAQDEQDAAAASVRE 171  
 QY 54 AEILSVLSHNIIOFYGVILEPNNYGVTEYASLSGLYVI-----NSNRSEE 101  
 DB 172 ARLPALMRHNIIEELRGVCLIQPHCLVLEFARGALNRLAANAAPPDPRAPGRARR 231  
 QY 102 MDMDHMTATDVAKGMHYLHMEAPVKVIRDLKSRNVI-----AADGVLCICDFG 153  
 DB 232 IPRHVLVNWAVQIARGLYLHBEAFVPIIRDLKSNILLEKIEHDI CNKTLKITDFG 291  
 QY 154 ASRFNHTHMSLVGTFPMMADEVIOQLPVSETCDTYSYGVVLMELTREVPFKLEGIO 213  
 DB 292 LAEMWRTTKMSTAGTYAMMADEVIKSILFSKSDIWSGYVLMELTGEVYRGIDGLA 351  
 QY 214 VAMLVYKERNLTIPSSCPSPFAELHOCWEADAKRPSFKOIIISLSMSND--TSLPD 271  
 DB 352 VAVGVAVNKLTPIPSTCEPFAKLMEKQWODPHIRPSFALILEQALVIGAVVTEMPO 411  
 QY 272 KCSNPLHNAKRECEIEATLERLKLERDLSEKQEL-----KERRRRLKMEOKLTE 324  
 DB 412 E--SFHSMQDWKLEIQWFDELRTKEKELRSREBELTRALLOQKSQBELKREQOLAE 469  
 QY 325 Q-----SNTPLLLPLAARSEESYFESKTEESNSA 354  
 DB 470 REIDVLERELNLIIFOLNQEKPVKRKGKFRSRRLKLDGHRISLPSPDFOKHITVQASP 529  
 QY 355 EM-----SCQITATSNGEHGNNPSLOAMML 380  
 DB 530 NLDKRRSLNSSSSSPSPPTMPLRALIQL 559

RESULT 12

AAB80923  
 ID AAB80923 standard; Protein; 1036 AA.

XX ABB80923;

XX 08-OCT-2002 (first entry)

DB Novel human protein (NHP) kinase.

XX Novel human protein; NHP; kinase; human; enzyme.

OS Homo sapiens.

XX

XX Key Location/Qualifiers

FT Misc-difference 925 /note="encoded by WGT"

XX

PN WO200255685-A2.

XX 18-JUL-2002.

XX 10-DEC-2001; 2001WO-US47606.

XX 11-DEC-2000; 2000US-254744P.

XX (LEXI-) LEXICON GENETICS INC.

XX Hu Y, Kleke JA, Donoho G;

XX WPI; 2002-566739/60.

XX N-PSDB; ABN86357, ABN86358.

PT Novel human kinase polynucleotide encoding a protein that shares

PT structural similarity with animal kinases for therapeutic, diagnostic

PS and pharmacogenomic applications -

PS Claim 1; Page 37-39; 41pp; English.

CC The invention relates to a novel human protein (NHP), kinase that shares

CC structural similarity with animal kinases. The kinase polynucleotides are

CC useful in therapeutic, diagnostic and pharmacogenomic applications and

CC for identifying compounds that modulate, i.e. act as agonists or

CC antagonists of the gene expression or gene product activity. The present

CC sequence represents the NHP kinase.

Sequence 1036 AA;

Query Match 25.2%; Score 599.5; DB 23; Length 1036;  
 Best Local Similarity 33.3%; Pred. No. 4,4e-45;

Matches 150; Conservative 75; Mismatches 146; Indels 79; Gaps 9;

QY 6 ASFVQIFEDDLOFFENCGGSGFSVYRAKWIISQKEVAAYKLIK-----IEKE 53  
 DB 114 SSVVHVAFERLELKELGAGFGQVYRATW--QGGEVAVKAAQDEQDAAAASVRE 171  
 QY 54 AEILSVLSHNIIOFYGVILEPNNYGVTEYASLSGLYVI-----NSNRSEE 101  
 DB 172 ARLPALMRHNIIEELRGVCLIQPHCLVLEFARGALNRLAANAAPPDPRAPGRARR 231  
 QY 102 MDMDHMTATDVAKGMHYLHMEAPVKVIRDLKSRNVI-----AADGVLCICDFG 153  
 DB 232 IPRHVLVNWAVQIARGLYLHBEAFVPIIRDLKSNILLEKIEHDI CNKTLKITDFG 291  
 QY 154 ASRFNHTHMSLVGTFPMMADEVIOQLPVSETCDTYSYGVVLMELTREVPFKLEGIO 213  
 DB 292 LAEMWRTTKMSTAGTYAMMADEVIKSILFSKSDIWSGYVLMELTGEVYRGIDGLA 351  
 QY 214 VAMLVYKERNLTIPSSCPSPFAELHOCWEADAKRPSFKOIIISLSMSND--TSLPD 271  
 DB 352 VAVGVAVNKLTPIPSTCEPFAKLMEKQWODPHIRPSFALILEQALVIGAVVTEMPO 411  
 QY 272 KCSNPLHNAKRECEIEATLERLKLERDLSEKQEL-----KERRRRLKMEOKLTE 324  
 DB 412 E--SFHSMQDWKLEIQWFDELRTKEKELRSREBELTRALLOQKSQBELKREQOLAE 469  
 QY 325 Q-----SNTPLLLPLAARSEESYFESKTEESNSA 354  
 DB 470 REIDVLERELNLIIFOLNQEKPVKRKGKFRSRRLKLDGHRISLPSPDFOKHITVQASP 529  
 QY 355 EM-----SCQITATSNGEHGNNPSLOAMML 380  
 DB 530 NLDKRRSLNSSSSSPSPPTMPLRALIQL 559

RESULT 13

AAB21717  
 ID AAB21717 standard; Protein; 1097 AA.

XX AAB21717;

XX

DT 16-JUL-2002 (first entry)  
 XX Human PKIN-12 protein.  
 XX Human; kinase; enzyme; PKIN-12 protein; immune system disorder; anaemia;  
 KW acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease;  
 KW asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease;  
 KW AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma;  
 KW leukaemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder;  
 KW Down's syndrome; gene therapy; protein therapy; cytostatic.  
 XX Homo sapiens.  
 OS  
 XX Key  
 FH Peptide  
 FT 1..17  
 FT /label=Signal\_peptide  
 FT 18..1097  
 FT /note="Mature human PKIN-12 protein"  
 FT 55..114  
 FT /note="SH3 domain"  
 FT 144..403  
 FT /note="Eukaryotic protein kinase domain"  
 FT 146..396  
 FT /note="Protein kinase domain"  
 FT 163..396  
 FT /note="Protein kinase domain"  
 FT 220..233  
 FT /note="Tyrosine kinase catalytic domain"  
 FT 258..276  
 FT /note="Tyrosine kinase catalytic domain"  
 FT 311..321  
 FT /note="Tyrosine kinase catalytic domain"  
 FT 330..352  
 FT /note="Tyrosine kinase catalytic domain"  
 FT 374..396  
 FT /note="Tyrosine kinase catalytic domain"  
 FT 438..749  
 FT /note="Leucine zipper domain"  
 FT 869..893  
 FT /note="Leucine zipper domain"  
 FT Domain  
 XX WO200218557-A2.  
 XX  
 PD 07-MAR-2002.  
 XX  
 PD 31-AUG-2001; 2001WO-US27219.  
 XX  
 PF 31-AUG-2000; 2000US-229873P.  
 XX 08-SEP-2000; 2000US-231357P.  
 XX 14-SEP-2000; 2000US-232654P.  
 PR 22-SEP-2000; 2000US-234902P.  
 PR 29-SEP-2000; 2000US-236499P.  
 PR 06-OCT-2000; 2000US-238389P.  
 PR 13-OCT-2000; 2000US-240542P.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX  
 PI Bandman O, Nguyen DB, Walla NK, Hafalia AJA, Yao MG, Gandhi AR,  
 PI Gururajan R, Ding L, Patterson C, Yue H, Baughn MR, Tribouley CM,  
 PI Thornton M, Elliott VS, Lu Y, Ison CH, Au-Young J, Tang YT,  
 PI Atimzai Y, Burdill JD, Marcus GA, Zingler KA, Lu DM, Lal PG,  
 PI Ramkumar J, Warren BA, Kearney L, Policky JL, Thangavelu K,  
 PI Burford N;  
 XX  
 DR WPI; 2002-329769/36.  
 DR N-PSDB; AAD34309.  
 XX  
 XX New human kinases, useful for diagnosing, treating or preventing immune  
 PT system disorders (e.g. Crohn's disease), neurological disorders (e.g.  
 PT epilepsy), or cell proliferative disorders (e.g. cancers such as  
 PT leukemia or lymphoma)  
 XX  
 PS Claim 67; Page 171-173; 218pp; English.

XX The present invention relates to human kinases (PKIN) and polynucleotides  
 CC encoding such proteins. PKIN sequences of the invention are useful for  
 CC diagnosing, treating or preventing disorders associated with aberrant  
 CC expression of PKIN, particularly immune system disorders (e.g. acquired  
 CC immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,  
 CC anaemia, asthma), neurological disorders (e.g. epilepsy, Charcot-Marie-  
 CC Tooth disease or seizures), cell proliferative disorders (e.g. cancers  
 CC such as adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma, sarcoma),  
 CC and developmental disorders (e.g. Down's syndrome). They are also used  
 CC in gene therapy and protein therapy. The present sequence is human  
 CC PKIN-12 protein.  
 XX  
 SQ Sequence 1097 AA;  
 XX  
 Query Match 24.8%; Score 590; DB 23; Length 1097;  
 Best Local Similarity 39.0%; Pred. No. 3,5e-44;  
 Matches 134; Conservative 60; Mismatches 118; Indels 32; Gaps 5;  
 QY 9 VQIKPDDLOFPENCGGSGSVYRAKWSODEKAVK-----KLLKIEKEAEI 56  
 DB 137 LEIDFAELTLEETIGGFGKRYRAFICD--EVAVKAARDPDSDISQTIENVROEAKL 194  
 QY 57 LSVLSHRNIIQFYGVLEPPNYGIVETASLGSLYDYNNSRSEMDPHMTATDVAK 116  
 DB 195 FAWLKHPNIIALRGVCLKPNCIWMFARGGPILRVSGRK--IPDILVWMAVQIAR 251  
 QY 117 GMAYLMEAPVKYIHRDLKSRNVV-----AADGVLCDFGASRPHNHTTMSLVG 168  
 DB 252 GMYVLDDEALVPIIHRDLKSSNLILOKVENGLSKIKIKIDFGIAEMWHTTMSAAG 311  
 QY 169 TFPMAPEVIOQLPVSFCDTYSYGVVLEMLTREVPFGELGLOVAMLVKNERLITP 228  
 DB 312 TYAMMAPEVIRAMFSGSDVMSYGVLLWELLTGEVFFRGIDGLAVAYVAMNKALPIP 371  
 QY 229 SSCPSFAELHOCWEADAKRPSFKQIISILESNDNDSLDKNSFHNKAEMRCEIE 288  
 DB 372 STCEPFAKLMEDCWNPDPSRPSFTNILDQLTIESGPFEMPKDSFICLDNMKHEIQ 431  
 QY 289 ATLERLKKLERDLSFKEOEL-----KERERLKNMEQKLTLEQ 325  
 DB 432 EMPDQRAKEKEKLTWEELTRAAQQKNEELRRRDEQIAER 475  
 XX  
 RESULT 14  
 AAE11775  
 ID AAE11775 standard; Protein; 1046 AA.  
 XX  
 AC AAE11775;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Human kinase (PKIN)-9 protein.  
 XX  
 XX Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;  
 KW cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;  
 KW AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;  
 KW atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;  
 KW myasthenia gravis; cirrhosis; cataract; growth and development disorder;  
 KW seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;  
 KW lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;  
 KW obesity; restorative therapy; immunomodulatory; vaccine; cardiovascular;  
 KW antimicrobial; cytostatic; antiinflammatory; asthma.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT 55..114  
 FT /note="SH3 domain"  
 FT 134..393  
 FT /note="Eukaryotic protein kinase domain"  
 FT 136..386  
 FT /note="Protein kinase domain"



FT Region 154..207  
 FT /note= "Receptor tyrosine kinase"  
 FT 181..228  
 FT /note= "Receptor tyrosine kinase"  
 FT Region 210..223  
 FT /note= "Tyrosine kinase catalytic site"  
 FT Region 232..254  
 FT /note= "Receptor tyrosine kinase"  
 FT 248..266  
 FT /note= "Tyrosine kinase catalytic site"  
 FT Region 290..337  
 FT /note= "Receptor tyrosine kinase"  
 FT Region 291..340  
 FT /note= "Receptor tyrosine kinase"  
 FT 298..330  
 FT /note= "Receptor tyrosine kinase"  
 FT Region 301..311  
 FT /note= "Tyrosine kinase catalytic site"  
 FT Region 320..342  
 FT /note= "Tyrosine kinase catalytic site"  
 FT 337..389  
 FT /note= "Receptor tyrosine kinase"  
 FT Region 345..389  
 FT /note= "Receptor tyrosine kinase"  
 FT 356..404  
 FT /note= "Receptor tyrosine kinase"  
 FT 364..386  
 FT /note= "Tyrosine kinase catalytic site"  
 FT  
 PN MO200181555-A2.  
 PD 01-NOV-2001.  
 PD  
 PF 20-APR-2001; 2001WO-US12992.  
 PF  
 PR 20-APR-2000; 2000US-199021P.  
 PR 20-APR-2000; 2000US-200226P.  
 PR 05-MAY-2000; 2000US-202339P.  
 PR 11-MAY-2000; 2000US-203505P.  
 PR 18-MAY-2000; 2000US-205564P.  
 PR 26-MAY-2000; 2000US-207739P.  
 PR 01-JUN-2000; 2000US-208795P.  
 PR  
 PA (INCY-) INCYTE GENOMICS INC.  
 PI Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB,  
 Bandman O, Lu DM, Lai P, Burford N, Khan PA, Walla NK, Yao MG,  
 Paterson C, Burrill JD, Marcus GA, Zingler KA, Reclapon SA, Lu Y,  
 Policky JL, Thornton M, Tang YT, Hatella A, Elliott VS, Baughn MR,  
 Gurrarajan R,  
 Gururajan R,  
 WPI; 2001-611740/70.  
 DR N-PSDB; MAD18824.  
 DR  
 XX  
 XX  
 FT Human kinases and nucleic acids, useful for preventing diagnosing and  
 FT treating cancers, inflammation and immune disorders -  
 PS  
 PS Claim 1; Page 134-136; 166pp; English.

CC supplement the patients own production of PKIN. PKIN nucleic acids may be  
 CC used to produce the PKIN polypeptide, by inserting the nucleic acids into  
 CC a host cell and culturing the cell to express the protein. PKIN nucleic  
 CC acid and its complementary sequences may also be used as DNA probes in  
 CC diagnostic assays to detect and quantitate the presence of similar  
 CC nucleic acid sequences in samples and therefore which patients may be  
 CC in need of restorative therapy. The present sequence is human PKIN-9  
 CC protein.  
 CC  
 XX  
 SQ Sequence 1046 AA.  
 Query Match 24.8%; Score 589; DB 22; Length 1046;  
 Best local similarity 39.1%; Pred. No. 4.1e-44;  
 Matches 134; Conservative 59; Mismatches 118; Indels 32; Gaps 5;  
 QY 10 QIKPDDLOFFENCSSGSGSYRAKMTISQKEVAVK-----KLTKEKEAEL 57  
 DB 128 EIDPRLTEETIIGIGFGKVRFAWIGD--EVAVKARHDPEDDISQTIENVROBKLP 185  
 QY 58 SVLSHRNIOFYGVILPPNYGVTEYASLSLYDYNRSRSEMDHMTWATDVAKG 117  
 DB 186 AMLKHPNIALRGVCLKEPMLCLMEFARCGPLNRVLSGR--IPDILVMVAQIARG 242  
 QY 118 MHYLMEAPVKVIRHDKSNVVI-----AADGLKICDGSARFPHNHTTMSLVGT 169  
 DB 243 MNYLDEAIVPIIHRDKSSNIIILQKVENGDLSNKLITPGLAREHRTTKSACT 302  
 QY 170 FPMWAPVIOSLPVESECDTYSYGVVLMELTRVFPKGLBGLQVAMLVKNERLTIPS 229  
 DB 303 YAMWAPVIRASMFSGDWSGVLMELTGEVFPFRIIDGLAAYGVANMKLALPIS 362  
 QY 230 SCPSFRELHQEWADAKKPSFKQIISILSMNSNDTSLPDKCNSFLHNKAWRCCEIA 289  
 DB 363 TCPEPFAKLMECDWMDPDRSPFTNIIDQLTTIBESGFEMPKDSFICLDNMWHEIOE 422  
 QY 290 TLERLKLKRDLSFKOEI-----KERERLKWMEOKLTBO 325  
 DB 423 MFDQLAKEKEKRLTWEBELTRALQCKOEELRRREDELAER 465  
 RESULT 15  
 AAE22763  
 ID AAE22763 standard; Protein; 847 AA.  
 XX  
 AC AAE22763;  
 XX  
 DT 09-AUG-2002 (first entry)  
 XX  
 DE Human mitogen activated protein kinase, MAP3K1.  
 XX  
 KW Human; cytosolic; antisense gene therapy; screening; protein kinase;  
 KW cancer; liver; colon; tumour; inflammation; arthritic synovium; MAP3K1;  
 XX  
 OS Homo sapiens.  
 XX  
 EN WO200224947-A2.  
 PD 28-MAR-2002.  
 PD  
 PF 20-SEP-2001; 2001WO-IB02237.  
 PF  
 PR 20-SEP-2000; 2000US-233999P.  
 PR 02-OCT-2000; 2000US-237419P.  
 PR 02-OCT-2000; 2000US-237423P.  
 PR 04-OCT-2000; 2000US-238558P.  
 PR 10-MAY-2001; 2001US-290555P.  
 XX  
 PA (KINE-) KINETEK PHARM INC.  
 PA (UYER-) UNIV BRITISH COLUMBIA.  
 XX  
 XX Yoganathan T, Delaney AD;  
 PI  
 XX

DR WPI: 2002-394145/42.  
DR N-PSDB; AAD6139.

XX Diagnosing cancer, comprises determining the upregulation of expression  
PT of a nucleic acid sequence encoding a protein kinase or upregulation of  
PT expression of the protein kinase, in the cancer

XX  
PS Claim 1; Page 60-62; 87p; English.

XX The invention relates to a method for screening biologically active agent  
CC that modulates cancer associated protein kinase function. The invention  
CC also relates to a method for diagnosing cancer comprising determining the  
CC upregulation of expression of a nucleic acid sequence encoding a protein  
CC kinase. The method is useful for diagnosing cancer. A protein kinase is  
CC useful for screening biological agents that modulate cancer associated  
CC protein kinase function. Downregulating the activity of protein kinase is  
CC useful for inhibiting the growth of a cancer cell, e.g. liver or colon  
CC cancer. A nucleic acid encoding protein kinase is useful to screen biopsy  
CC derived tumours and inflammatory samples such as arthritic synovium, for  
CC amplified DNA in the cell or increased expression of corresponding mRNA  
CC or protein and is also useful to detect differences in expression levels  
CC such as molecular weight, amino acid and nucleotide sequences between the  
CC two cells. The present sequence is human mitogen activated protein  
CC kinase, MAP3K11.

CC  
XX  
SQ Sequence 847 AA;

Query Match

Best Local Similarity 24.5%; Score 581.5; DB 23; Length 847;  
Matches 132; Conservative 37.4%; Pred. No. 1,4e-43; Indels 39; Gaps 7;  
Matches 132; Conservative 60; Mismatches 122;

OY 13 FDLQFFPENGCGSGFSGVYRAKWIISO-----DKENAVKKLLKIEKEAIIISVLS 61  
DB 114 FOELRLLEVYIGIGFGKTVRGSGRGLVAVKARODPEDISV-TAESVROEARLFAMLA 172  
OY 62 HRNIIQFYGVILEPPNYGIVTEYASLSGLYDYNRSREMDMDHM-TWATDVAKMHY 120  
DB 173 HPNIIALKAVCLEEPNLCIWEYAAAGPLSRALARRVP---PHLVNMAVOIARGMHY 228  
OY 121 LHMFAPIKVIHRDLKSNV-----IAADGV-----LKICDGSARFPHNHTTMSLVGTFPW 172  
DB 229 LHCEALVPVHRDLKSNVILLOPIESDMEHNTLKTDFGLAREWHKTTOMSAAGTVAM 288  
OY 173 MAPVIOISLPVSECTDYSGVIVEMTLTREVPKGLBGLQVAMLVVEKNERLTIPSSCP 232  
DB 289 MAPEVTKASTFGSGDVMSFGVILMELLTGVPYRGIDCLAVAGVAVNKLTPIPSTCP 348  
OY 233 RSPFALLHQCEWADAKKRPSPFKOIIISLESMSNDTSLPDKNSFLHNKAEWRCEIATLE 292  
DB 349 EPEAQLMADCAWQDPHRRPASTILOQLEALEAQLREMPRDSFHSWQEGMKREIOGLPD 408  
OY 293 RLKRLERDLSPKEOE-----LKERERLKMWEOKLTQOSNTPLL 331  
DB 409 ELRAREKELLSEBELTRAREORSQAEOLRRREHLAQMELVEFERELTLL 461

Search completed: May 1, 2003, 20:35:33  
Job time : 92.5732 secs

GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using bw model

Run on: May 1, 2003, 20:38:08 ; Search time 29.1911 Seconds  
(without alignments)  
1344.947 Million cell updates/sec

Title: US-09-757-982-5  
Perfect score: 2375  
Sequence: 1 MSSLGASFWQIKFDDLOFFE.....GDDDDDDGEEEDNDNDNSE 455

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328255 seqs, 86286685 residues  
Total number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2375	100.0	455	US-09-757-982-5	Sequence 5, Appl1
2	599.5	25.2	1036	US-10-014-882-2	Sequence 2, Appl1
3	593	25.0	394	US-09-862-027-19	Sequence 19, Appl1
4	581.5	24.5	847	US-10-143-133-2	Sequence 2, Appl1
5	568.5	23.9	966	US-09-771-161A-197	Sequence 197, App
6	545.5	23.0	328	US-09-862-027-18	Sequence 18, Appl1
7	504.5	21.2	850	US-09-904-389-2	Sequence 2, Appl1
8	487	20.5	263	US-09-840-704-5	Sequence 5, Appl1
9	481	20.3	579	US-10-158-895-4	Sequence 4, Appl1
10	481	20.3	590	US-10-158-895-15	Sequence 15, Appl1
11	424	17.9	92	US-09-764-868-799	Sequence 799, App
12	392.5	16.5	835	US-09-947-199-8	Sequence 8, Appl1
13	385	16.2	1036	US-09-771-161A-255	Sequence 255, App
14	385	16.2	1036	US-09-771-161A-256	Sequence 256, App
15	382.5	16.1	251	US-08-987-689A-32	Sequence 32, Appl1
16	379.5	16.0	505	US-09-977-260-6	Sequence 6, Appl1
17	379.5	16.0	505	US-09-977-261-6	Sequence 6, Appl1
18	379.5	16.0	505	US-09-977-269-6	Sequence 6, Appl1
19	379.5	16.0	505	US-09-982-610-20	Sequence 20, Appl1

20	377	15.9	835	10	US-09-947-199-2	Sequence 2, Appl1
21	376.5	15.9	537	9	US-09-977-260-11	Sequence 11, Appl1
22	376.5	15.9	537	9	US-09-977-261-11	Sequence 11, Appl1
23	376.5	15.9	537	10	US-09-977-269-11	Sequence 11, Appl1
24	374.5	15.8	537	10	US-09-771-161A-212	Sequence 212, App
25	374.5	15.8	537	10	US-09-771-161A-213	Sequence 213, App
26	365.5	15.4	536	9	US-09-977-260-12	Sequence 12, Appl1
27	365.5	15.4	536	9	US-09-977-261-12	Sequence 12, Appl1
28	365.5	15.4	536	10	US-09-977-269-12	Sequence 12, Appl1
29	358.5	15.1	675	9	US-10-186-399-3	Sequence 3, Appl1
30	358.5	15.1	675	9	US-09-977-260-4	Sequence 4, Appl1
31	358.5	15.1	675	9	US-09-977-261-4	Sequence 4, Appl1
32	358.5	15.1	675	10	US-09-977-269-4	Sequence 4, Appl1
33	358	15.1	764	10	US-09-925-302-714	Sequence 714, App
34	357.5	15.1	425	10	US-09-828-313-29	Sequence 29, Appl1
35	355	14.9	310	9	US-09-939-833-7	Sequence 7, Appl1
36	355	14.9	310	10	US-09-939-834-7	Sequence 7, Appl1
37	355	14.9	310	10	US-09-939-832-7	Sequence 7, Appl1
38	355	14.9	822	9	US-09-757-415A-2	Sequence 2, Appl1
39	349.5	14.7	536	9	US-09-977-260-13	Sequence 13, Appl1
40	349.5	14.7	536	9	US-09-929-266-10	Sequence 10, Appl1
41	349.5	14.7	536	9	US-09-977-261-13	Sequence 13, Appl1
42	349.5	14.7	536	10	US-09-977-269-13	Sequence 13, Appl1
43	349.5	14.7	1308	10	US-09-940-101-2	Sequence 2, Appl1
44	347.5	14.6	450	9	US-09-977-260-7	Sequence 7, Appl1
45	347.5	14.6	450	9	US-09-977-261-7	Sequence 7, Appl1

## ALIGNMENTS

RESULT 1	US-09-757-982-5	Application US/09757982
Sequence 5	Application US/09757982	
Patent No.	US20020094559A1	
GENERAL INFORMATION:		
APPLICANT:	Accon, Susan	
TITLE OF INVENTION:	NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR	
FILE REFERENCE:	NMI-050	
CURRENT APPLICATION NUMBER:	US/09/757, 982	
CURRENT FILING DATE:	2001-01-10	
PRIOR APPLICATION NUMBER:	09/163, 115	
PRIOR FILING DATE:	1998-09-29	
NUMBER OF SEQ ID NOS:	15	
SOFTWARE:	PatentIn Ver. 2.0	
SEQ ID NO 5		
LENGTH:	455	
TYPE:	PRT	
ORGANISM:	Homo sapiens	
US-09-757-982-5		
Query Match	100.0%; Score 2375; DB 10; Length 455;	
Best Local Similarity	100.0%; Pred. No. 1.7e-168;	
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	MSSLGASFWQIKFDDLOFFENCSSGSGSVYRAKMIQDKEVAVAKLTKIEKEAIIISVL 60
DB	1	MSSLGASFWQIKFDDLOFFENCSSGSGSVYRAKMIQDKEVAVAKLTKIEKEAIIISVL 60
QY	61	SHRNIIQFYGVILPEPNYGVITVEYASLSGLYDINSNRSEEMDMHIMTAADVAKGMHY 120
DB	61	SHRNIIQFYGVILPEPNYGVITVEYASLSGLYDINSNRSEEMDMHIMTAADVAKGMHY 120
QY	121	LHMAAPVKVYHRDLKSNVVIADGVYKICDPFGASRPHNHTHMSLVGTFFMAPEVIOS 180
DB	121	LHMAAPVKVYHRDLKSNVVIADGVYKICDPFGASRPHNHTHMSLVGTFFMAPEVIOS 180
QY	181	LPVSETCDTYSYGVVLEMLTREVPFGKLGLOVAMVVEKNRELTTPSSCPSPFAELIH 240
DB	181	LPVSETCDTYSYGVVLEMLTREVPFGKLGLOVAMVVEKNRELTTPSSCPSPFAELIH 240
QY	241	QCMBAKKKPPSPFOIISILSMSNDTSLPDKCNSFLHNAWRCIEATLERLKJLERD 300

Db 241 QWMEADAKRPSFKQIISILESMNDISLDPKCNSEFLHNKAEWCEIATLERLKLERD 300  
Qy 301 LSPFEQELKEBERLTKWMEQKLTROSNTPLLPLAARMSSESYFSPKTEESNAEMSCQI 360  
Db 301 LSPFEQELKEBERLTKWMEQKLTROSNTPLLPLAARMSSESYFSPKTEESNAEMSCQI 360  
Qy 361 TATNGEGHGNPNLSLOAMLMGFGDIFSMNRAGAVMHSGMOINMOAKONSKTTSKRKGK 420  
Db 361 TATNGEGHGNPNLSLOAMLMGFGDIFSMNRAGAVMHSGMOINMOAKONSKTTSKRKGK 420  
Qy 421 KVMNALGSPDLSSEGDDEDDDDGEEEDNDNDNSE 455  
Db 421 KVMNALGSPDLSSEGDDEDDDDGEEEDNDNDNSE 455

## RESULT 2

US-10-014-882-2  
Sequence 2, Application US/10014882  
Patent No. US20020107384A1  
GENERAL INFORMATION:

APPLICANT: Hu, Yi  
APPLICANT: Kieke, James  
APPLICANT: Donoho, Gregory  
TITLE OF INVENTION: No. US20020107384A1 Human Kinase and Polynucleotides Encoding  
FILE REFERENCE: LEX-0279-USA  
CURRENT APPLICATION NUMBER: US/10/014,882  
PRIOR FILING DATE: 2001-12-11  
PRIOR APPLICATION NUMBER: US 60/254,744  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 1036  
TYPE: PRT  
ORGANISM: homo sapiens  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)...(1036)  
OTHER INFORMATION: Xaa = Any Amino Acid  
US-10-014-882-2

Query Match 25.2%; Score 599.5; DB 12; Length 1036;  
Best Local Similarity 33.3%; Pred. No. 2.6e-36;  
Matches 150; Conservative 75; Mismatches 146; Indels 79; Gaps 9;

Qy 6 ASFVQIKRDDI OFPENCSSGSGSYRAKWTISQKEVAVKLLK-----IKKE 53  
Db 114 SSPVHAFERLELKEILGAGGQVYRAWM--QGGEVAVKARODPEDAAAAEVSARE 171  
Db 54 AEILSVSHRNIIQGYVILEPPNYGIYEVASLSGLDYI-----NSNRSEB 101  
Db 172 ARLRAMLHPHIIELRGVCLQPFILVLFPRAGALRALAANAADPPRAGRARR 231  
Qy 102 MDMDHMTWATDVAKGMHYLHMEAFVYIHRDLKSRNVI-----AADGVLCIDFG 153  
Db 232 IPRVILVMAVAQIARGMVYLHBEAFVPIIHRDLKSNITLLEKTEHDDICNKTILKITDFG 291  
Qy 154 ASRPHNHTTHSLVGTFFPMAAPVYQSLPVSETCTTYSYGVLMEMLTREVPFKLEGLQ 213  
Db 292 LARSMHRTTKKSTAGTYAMAPVYKSLFSGSDIMSYGLVWLLTGEVPRGIDGLA 351  
Qy 214 VAMLVENKNEBLTTPSGCPSPFALLHQWMEADAKRPSFKQIISILESMND--TSLPD 271  
Db 352 VAIGVAVNKLTLPISTCPPEPPAKLMKECWODPHIRSPALLIQLVLAIEGAVTWPQ 411  
Qy 272 KCNSEFLHNKAEWCEIATLERLKLERDLSFKEQEL-----KERERLTKWMEQKLT 324  
Db 412 E--SFHSMQDDMKLEIQMPBELRTKEKELNSREBELRALAQSGEELKXRRQGLAF 469  
Qy 325 Q-----SNTPLLLPLAARMSSESYFSPKTEESNA 354  
Db 470 REIDVLERELNILLFQLNQEKPKVKKRKGKFKRSRLKLDGHRISLPSDFGHKITVQASP 529

Qy 355 EM-----SCQITATNGEGHGNPNLSLOAMLT 380  
Db 530 NLDKRRSLNSSSSSPSSPTMPPRLRAIQL 559

## RESULT 3

US-09-862-027-19  
Sequence 19, Application US/09862027  
Patent No. US20020142428A1  
GENERAL INFORMATION:

APPLICANT: Hodge, Martin R.  
TITLE OF INVENTION: No. US20020142428A1 Kinases and Uses Thereof  
FILE REFERENCE: 35800/234862  
CURRENT APPLICATION NUMBER: US/09/862,027  
PRIOR FILING DATE: 2001-05-21  
PRIOR APPLICATION NUMBER: US 09/345,473  
NUMBER OF SEQ ID NOS: 82  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 19  
LENGTH: 394  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-862-027-19

Query Match 25.0%; Score 593; DB 10; Length 394;  
Best Local Similarity 39.3%; Pred. No. 2.4e-36;  
Matches 133; Conservative 59; Mismatches 114; Indels 32; Gaps 5;

Qy 15 DLOFPENCSSGSGSYRAKWTISQKEVAVK-----KLEKIEKAEILSVLSH 62  
Db 2 ELTLEELIIGGFGKVRARAFICD--EVAVAAARHDPDEDISOTIENVRQEAFLMLKH 59  
Qy 63 RNIQFVGLLEPPNYGIYEVASLSGLDYINSEBMDHMTWATDVAKGMHYL 122  
Db 60 PNIIALRGVCLKEPNLCLVMEFARGGPLNRYLSGR--IPPDILVMAVAQIARGMVYLH 116  
Qy 123 MEAFVYIHRDLKSRNVI-----AADGVLCIDFGASRFNHTTHSLVGTFFPMA 174  
Db 117 DEAIVPPIIHRDLKSNITLILQKVENGLSKILKITPGLAREBHRTTKNSAAGTYAMA 176  
Qy 175 PEVIQSLPVSETCTTYSYGVLMEMLTREVPFKLEGLQVAMLVENKNEBLTTPSCPRS 234  
Db 177 PEVIRASMFSGSDIMSYGLVWLLTGEVPRGIDGLRAVYGVAMNKLALPISTCPEP 236  
Qy 235 FAEILHQWMEADAKRPSFKQIISILESMNDISLDPKCNSEFLHNKAEWCEIATLERL 294  
Db 237 FAKLMEQCMNPDPSRPSFTNILDQLTTIBESGFEMPXDSFCLQDNWGHIEQMFQDL 296  
Qy 295 KLEERDLSFKEQEL-----KERERLTKWMEQKLTREQ 325  
Db 297 RAKEKEIETWEELTRALAQKNOBELRRREQGLAR 334

## RESULT 4

US-10-143-133-2  
Sequence 2, Application US/10143133  
Patent No. US20020197658A1  
GENERAL INFORMATION:

APPLICANT: Delaney, Allen  
APPLICANT: Yoganathan, Thillainathan  
TITLE OF INVENTION: Cancer Associated Protein Kinase and Its Use  
FILE REFERENCE: KINE-023  
CURRENT APPLICATION NUMBER: US/10/143,133  
PRIOR FILING DATE: 2002-05-09  
PRIOR APPLICATION NUMBER: 60/290,555  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 847  
TYPE: PRT  
ORGANISM: Homo sapien

US-10-143-433-2

Query Match 24.5%; Score 581.5; DB 9; Length 847;

Best Local Similarity 37.4%; Pred. No. 4,4e-35; Indels 39; Gaps 7;

Matches 132; Conservative 60; Mismatches 122; Indels 39; Gaps 7;

13 FDDLOFFENCSSGSGSYRAKMSIQ-----DKEVAVKLLKIEKEAELISVLS 61

114 FOELRELEVICIGGKRYRGSGMGLVAVKARQDDEDISV-TASVGEARLFLPMLA 172

62 HRNIIQFVILPEPPNYGIVTEYASLSGYDINSNSEMDMDHIM-TATVAGKMY 120

173 HPIIILAKAVLEEPNCLVMEYAGPLRAGRRVP---PHVLNVAVOIAGKMY 228

121 LHMZAPKVIHRDLKSNV---IAADGV---LKIICPGASRFHNTHTMSLVGTFPM 172

229 LHCEALPVVHRDLKSNVILLOPTESDOMEHKLITDGLAREMHKTTQMSAATTYAM 288

173 MAPEVIOSLPSETCDTVSYGVVLMEMLTREVPFKGLEQVAMLVYKNERLTIPSSCP 232

289 MAPEVIOASTPSSKSDVMSFGLWELTGEVPRGIDCLAVGVAVVVKLTLPSTCP 348

233 RSFAELHQCMEADAKKPSFKOIIISLESMSNDTSLPDKNSFLNKAEMCEIATLE 292

349 EPPAQLMADCMADPHRRPDFASILOLELAQVLEMPRDSFHMOSGCKREIQLFPD 408

293 RLKLEKRLSFKOE-----LKERERLKMEOKLTQSNTPPL 331

409 ELBAKEKELLSREBELTRARERQSAEQLRREHLLAQMELEVERELTLL 461

RESULT 5

US-09-771-161A-197

Sequence 197 Application US/09771161A

Patent No. US20020110811A1

GENERAL INFORMATION:

APPLICANT: LEVINE, et al.

TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES

FILE REFERENCE: 802620-2005.1

CURRENT APPLICATION NUMBER: US/09/771,161A

CURRENT FILING DATE: 2001-01-26

PRIOR FILING DATE: 09/724,676

PRIOR APPLICATION NUMBER: 136776

PRIOR FILING DATE: 2000-06-15

PRIOR APPLICATION NUMBER: 135619

PRIOR FILING DATE: 2000-04-12

NUMBER OF SEQ ID NOS: 273

SOFTWARE: Patent version 3.0

SEQ ID NO 197

LENGTH: 966

TYPE: PRT

ORGANISM: Homo sapiens

US-09-771-161A-197

Query Match 23.9%; Score 568.5; DB 10; Length 966;

Best Local Similarity 31.4%; Pred. No. 4,8e-33; Indels 61; Gaps 13;

Matches 145; Conservative 89; Mismatches 167; Indels 61; Gaps 13;

10 OIKFPDLOFFENCSSGSGSYRAKMSIQDEKAVKLLKIEKEAELT-SVLSHNNIIQ 67

162 EYFEEISELQWLSGAGAVFLGFP--RAEEVAIKK-VREONETDIKHLKRLGHPNIIA 218

68 FYGVILPEPPNYGIVTEYASLSGYDINSNSEMDMDHIM-TATVAGKMYHMAFV 127

219 FKGVCTAPCYCIIMEYCAHOLYEVLAGR--KITRLVDMSTGIAAGNYYHLH-- 273

128 KVIRDLKSNVVAIAADGVKICPGASR-FHNTHTMSLVGTFPMMAPEVIOSLPVSET 186

274 KIHRLKSNVVAIAADGVKICPGASR-FHNTHTMSLVGTFPMMAPEVIOSLPVSET 186

187 CDTYSYGVVLMEMLTREVPFKGLEQVAMLVYKNERLTIPSSCPSEFELLHQCMEAD 246

DB 334 VDISFVVLWELTGEIPYDVDSALIWGVGNSLHLVPSTCPDGFILMKQTWQSK 393

QY 247 AKRPSFKOIIISLESMSNDTSLPDKNSFLNKAEMCEIATLETLKLERLDSFKQ 306

DB 394 PNRPSFRQTLMLHLDIASADV-LATPOETFKSQAEKREYKHFEXIKSEGTIHLDE 452

QY 307 ELKERER-----LKMEOKLTQSNTPPLPLAARMSSESYFESKTEESNAEMSC 358

DB 453 ELIRREELRLHALDIRHYERKLERANN--LYMELSAIMIQ--LEMRKELIKREQAV 507

QY 359 QITATSNSEGHGNPSL--QAMLMGFGDIFSNMKAGAVHSGWOI----- 402

DB 508 EKKYPTGYKRPVPIIHPNMEKL-----MKRQVPHKSGKQTRPDLRLSEGIPTT 560

QY 403 -----NMOAKONSSTTSRRKQKVMALGFSDF 431

DB 561 EVAPASPLSGSPKMSSTSSKSRYSKRRRRGNSRQSHDF 602

RESULT 6

US-09-862-027-18

Sequence 18 Application US/09862027

Patent No. US20020142428A1

GENERAL INFORMATION:

APPLICANT: Hodge, Martin R.

TITLE OF INVENTION: No. US20020142428A1el Kinases and Uses Thereof

FILE REFERENCE: 35800/234862

CURRENT APPLICATION NUMBER: US/09/862,027

CURRENT FILING DATE: 2001-05-21

PRIOR FILING DATE: 09/345,473

PRIOR APPLICATION NUMBER: 1999-06-30

NUMBER OF SEQ ID NOS: 82

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 18

LENGTH: 328

TYPE: PRT

ORGANISM: C. elegans

US-09-862-027-18

Query Match 23.0%; Score 545.5; DB 10; Length 328;

Best Local Similarity 41.2%; Pred. No. 6,4e-33; Indels 13; Gaps 5;

Matches 112; Conservative 54; Mismatches 93; Indels 13; Gaps 5;

QY 8 FVOIKEDLOFFENCSSGSGSYRAKMSIQD---KEVAVKLLKIEKEAELTSLSHRN 64

DB 44 FPDQRDIDQVGDHIGVGFAGVPSGNWTLPGSRTIALKKVFLERAEILSKIRHN 103

QY 65 IIOFYVILPEPPNYGIVTEYASLSGYDINSNSEEM-----DMDHIMTATDVAK 116

DB 104 IIOFYVICATGNDFFIYVEYAEKSLYDFIHSESSQSFASSGNSPDDVYKMASQIAS 163

QY 117 GMYHLEAPVAVIHRDLKSNVVAIAADGVKICPGASRFHNT-TMSLVGTFPMNAP 175

DB 164 GIQYHLDVAVTIIHRDLKSNVVAIAADGVKICPGASRFHNT-TMSLVGTFPMNAP 223

QY 176 E-VIOSLPVSETCDTVSYGVVLMEMLTREVPFKGLEQVAMLVYKNERLTIPSSCP 234

DB 224 EMILQSEGLTATDVMSYGVVLMELISKEVYKQYSESRITMTTQSSITLALPSCAP 283

QY 235 FAEILHQCMEADAKKPSFKOIIISLESMSND 266

DB 284 LKQALMSNCMTPKDRAAMROIQELNRLNAGN 315

RESULT 7

US-09-904-389-2

Sequence 2 Application US/09904389

Patent No. US20020129404A1

GENERAL INFORMATION:

APPLICANT: Clendenen, Stephanie K.

TITLE OF INVENTION: CTR1 HOMOLOGUE FROM MELON

FILE REFERENCE: 4257-0029.30

CURRENT APPLICATION NUMBER: US/09/904,389  
CURRENT FILING DATE: 2001-07-12  
PRIOR APPLICATION NUMBER: US 60/218,307  
PRIOR FILING DATE: 2000-07-14  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 850  
TYPE: PRT  
ORGANISM: Cucumis melo  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (154) ..(154)  
OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-904-389-2

Query Match 21.2%; Score 504.5; DB 10; Length 850;  
Best Local Similarity 40.9%; Pred. No. 2,3e-29;  
Matches 112; Conservative 41; Mismatches 106; Indels 15; Gaps 4;

3 SLASFOVKEDDLOPFENCGGSGSVYRAKWTISODKEVAVKKLKIE-----K 52  
DB 564 SLGSDLVIPWTDLDRKRTAGSGFVYRGSHGSD--VAVKLTEDDFPERVNEFLR 621  
QY 53 EAEIISVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLDYDINSRSEMDMHIMTWAT 112  
DB 622 EVAIKSKLRHPVIVFMGAVTPEPMLSTVTEYLSRGLSLYRLHKSGVGDIDETRRINNAF 681  
QY 113 DVAKGMHYIAMEAPVKVYIHRDLKSNVVIADGVLIKIDFGASRPHNNT--THMSLVGT 170  
DB 682 DVAKGMNHYIHRDP--PIVHRDLKSNPLVDKTYKVCDFGLSRKANTPLSSSAATP 740  
QY 171 PMMAPEVIOSLPVSETCDYSGYVLMEMLTREVPFKGLEQVAMLVENKERTLPSS 230  
DB 741 EMMAPEVILDEPSNKSVDYSGVILMELATLQCPWCMNPAQVAAVGFGRKRLIDPRD 800  
QY 231 CPRFAELHQCMEADAKRPSFKOIIISLESMS 264  
DB 801 VNPKLASLIVACWADPEPKRPSFSSIMETLKPMT 834

RESULT 8  
US-09-840-704-5  
Sequence 5, Application US/09840704  
Patent No. US20020122801A1  
GENERAL INFORMATION:  
APPLICANT: Dedhar, Shoukat  
APPLICANT: Hamigan, Greg  
TITLE OF INVENTION: Intergrin-Linked Kinase and its Uses  
FILE REFERENCE: KIN-2CON  
CURRENT APPLICATION NUMBER: US/09/840,704  
CURRENT FILING DATE: 2001-04-23  
PRIOR APPLICATION NUMBER: 09/566,906  
PRIOR FILING DATE: 2000-05-09  
PRIOR APPLICATION NUMBER: US06/752,345  
PRIOR FILING DATE: 1996-11-19  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5  
LENGTH: 263  
TYPE: PRT  
ORGANISM: H. sapiens  
FEATURE:  
NAME/KEY: Other  
LOCATION: (1) ... (263)  
US-09-840-704-5

Query Match 20.5%; Score 487; DB 10; Length 263;  
Best Local Similarity 41.0%; Pred. No. 1.1e-28;  
Matches 109; Conservative 43; Mismatches 98; Indels 16; Gaps 5;

QY 11 IKFDLOPFENCGGSGSVYRAKWTISODKEVAVKKLKIE-----KEAEIISVL 60  
DB 11 IKFDLOPFENCGGSGSVYRAKWTISODKEVAVKKLKIE-----KEAEIISVL 60

DB 1 IPWCDLNIKEIKAGSGFCTYHRAEMHSGD--VAVKLTEDDFPERVNEFLR 58  
QY 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLDYDIN-SRSEMDMHIMTWATDVAKGM 119  
DB 59 RHPNIVLFMGAVTPEPMLSTVTEYLSRGLSLYRLHKSGAEQDERRRLSMAYDVAKGM 118  
QY 120 YLHMAPEVKVYIHRDLKSNVVIADGVLIKIDFGASRPHNNT--THMSLVGTPEWMAPEV 177  
DB 119 YLHNNP--PIVHRDLKSNPLVDKTYKVCDFGLSLKASTPLSSSAATPEWMAPEV 177  
QY 178 IQSLPVSETCDYSGYVLMEMLTREVPFKGLEQVAMLVENKERTLPSSCPRFAE 237  
DB 178 LDEPSNKSVDYSGVILMELATLQCPWCMNPAQVAAVGFGRKRLIDPRD 237  
QY 238 LIHQCEADAKRPSFKOIIISLESMS 263  
DB 238 IIEGCTNEPWRKPSFATIMDLRPL 263

RESULT 9  
US-10-158-895-4  
Sequence 4, Application US/10158895  
Patent No. US20020155624A1  
GENERAL INFORMATION:  
APPLICANT: ONO, KOICHIRO  
APPLICANT: TSUCHIYA, MASAYUKI  
APPLICANT: OHTOMO, TOSHIHIKO  
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
FILE REFERENCE: 053466/0278  
CURRENT APPLICATION NUMBER: US/10/158,895  
CURRENT FILING DATE: 2002-06-03  
PRIOR APPLICATION NUMBER: US/09/529,279  
PRIOR FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: PCT/JP98/04796  
PRIOR FILING DATE: 1998-10-22  
PRIOR APPLICATION NUMBER: JP 9/290188  
PRIOR FILING DATE: 1997-10-22  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 579  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-158-895-4

Query Match 20.3%; Score 481; DB 9; Length 579;  
Best Local Similarity 30.4%; Pred. No. 7.9e-28;  
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;

QY 7 SFVOIKFDDLOPFENCGGSGSVYRAKWTISODKEVAVKKLKIEKAE-----IL 57  
DB 27 NFEIIDEKIEIEVEVVGAFVVCRAK--RAKDVAIK--OIESESRKAFIVELRQL 81  
QY 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLDYDINSRSEMDMHIMTWATDV 114  
DB 82 SRVHPNIVLYKGCINP--VCLVMEYAEBSGLYNVLHG--APLPYVTAHMSWCLQC 137  
QY 115 AKGMHYIAMEAPVKVYIHRDLKSNVVIADGVLIKIDFGS--ASRFNHTHMSLVGFPW 172  
DB 138 SQGVAYLHSMQPKALIHRLDKSNPLVDKTYKVCDFGLACDIQTHMTNKK--GSAW 195  
QY 173 MAPEVIOSLPVSETCDYSGYVLMEMLTREVPFKGLEQ--LOVAMLVENKERTLPSS 230  
DB 196 MAPEVBSGNSYSEKODVPSGIILMEVITRRKPPDELIGGAPRIMW-AVHNGTRPPLIKN 254  
QY 231 CPRFAELHQCMEADAKRPSFKOIIISL-----ESMSNDT- 267  
DB 255 LKPIESLMTRCWSKDPQPSMEIIVKIMTHMRYPGADPELPQYQYSDGQNSAT 314  
QY 268 ---SLPDKNSLHKAWRCE-IEATLERLKLKRLDSFKEOLKREERRLKMEOKLT 323  
DB 315 STGSEMDIASITVTSKSDTNMEQVPAINDTITKRLSKLKNQAQCSGSLSL---GAS 371

QY 324 EQSNTPLLLPLAAMSEESYFESTERSNGAENSQOITATSNGEGHGMNLSIQAMLMGF 383  
DB 372 HGSSVESLPTPS-----EGKRMGADMSIEIARIAATT-GNGQRRRSIQDLIVTGT 421  
QY 384 --GDIFSMNKAQAV 395  
DB 422 EPGGVSSRSSSPSV 435

RESULT 10  
US-10-158-895-15  
Sequence 15, Application US/10158895  
Patent No. US20020155624A1  
GENERAL INFORMATION:  
APPLICANT: ONO, KOICHIRO  
APPLICANT: OHTOMO, TOSHIMIKO  
APPLICANT: TSUCHIYA, MASAYUKI  
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
FILE REFERENCE: 053466/0278  
CURRENT APPLICATION NUMBER: US/10/158,895  
PRIOR FILING DATE: 2002-06-03  
PRIOR APPLICATION NUMBER: US/09/529,279  
PRIOR FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: PCT/JP98/04796  
PRIOR FILING DATE: 1998-10-22  
PRIOR APPLICATION NUMBER: JP 9/290188  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 15  
LENGTH: 590  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-158-895-15

Query Match 20.3%; Score 481; DB 9; Length 590;  
Best Local Similarity 30.4%; Pred. No. 8.1e-28;  
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;

QY 7 SFVOIKFDDLOFFENCGGSPFSVYRAKVISODKEVAVKLIKIEKAE-----IL 57  
DB 27 NNEEDYVEIEVEEVGAGFVGVCAKWK--RAKDVAIK--QIESESEKAFIVELRQL 81  
QY 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDINSNSEEM---DMHIMTWADV 114  
DB 82 SRVNHPIVLYAGACLN--VCLVMEYAEGLSVINVLHG--AELPYTTAAHMSWCLQC 137  
QY 115 AKGMHYLMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRPHNHTTMSLVGTFPM 172  
DB 138 SOGVAVYLSHSMOPKALIHRLKPPNLLVAGGVVLKICDFGACDIQTHMTNKK--GSAAM 195  
QY 173 MAPEVIOSLPVSETCDTYSYGVVLEMLTREVPPKGLEG--LOVAVLVENKNERLTPSS 230  
DB 196 MAPEVIESNTSEKCDVSWGIIIMEVITRRKPFDEIGGPAFRIMV-AVHNGTTPPLIKN 254  
QY 231 CPRSFALLHOCWEADAKRPSFKOIIISL-----ESMSNDT- 267  
DB 255 LKPIESIMTRCWSKDPQSRSMIEIVIMTHLMRYPPGADEPLQYCOQSDSOGNSAT 314  
QY 268 ---SLPDKCNFPLNKAKWRCE-IEATLERLKLJERLDSFEKEQLKERERLKKMEQKLT 323  
DB 315 STGSFMDIATSTNKSSTJNMEQVPATNDITIKRLSKLKNQAQSGESGLSL--GAS 371  
QY 324 EQSNTPLLLPLAAMSEESYFESTERSNGAENSQOITATSNGEGHGMNLSIQAMLMGF 383  
DB 372 HGSSVESLPTPS-----EGKRMGADMSIEIARIAATT-GNGQRRRSIQDLIVTGT 421  
QY 384 --GDIFSMNKAQAV 395  
DB 422 EPGGVSSRSSSPSV 435

RESULT 11

US-09-764-868-799  
Sequence 799, Application US/09764868  
Patent No. US20020168711A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PT232  
CURRENT APPLICATION NUMBER: US/09/764,868  
PRIOR FILING DATE: 2001-01-17  
PRIOR APPLICATION data removed - refer to PALM or file wrapper  
NUMBER OF SEQ ID NOS: 1510  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 799  
LENGTH: 92  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (4)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-868-799

Query Match 17.9%; Score 424; DB 9; Length 92;  
Best Local Similarity 98.8%; Pred. No. 1.4e-24;  
Matches 80; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 53 EARLSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDINSNSEEMDMHIMTWAT 112  
DB 12 QAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDINSNSEEMDMHIMTWAT 71  
QY 113 DVAKGMHYLMEAPVKVIHRD 133  
DB 72 DVAKGMHYLMEAPVKVIHRD 92

RESULT 12  
US-09-947-199-8  
Sequence 8, Application US/09947199  
Patent No. US20020127684A1  
GENERAL INFORMATION:  
APPLICANT: Raju, Jayaseelan  
TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES  
FILE REFERENCE: MNI-068CP2  
CURRENT APPLICATION NUMBER: US/09/947,199  
PRIOR FILING DATE: 2001-09-05  
PRIOR APPLICATION NUMBER: 60/111,938  
PRIOR FILING DATE: 1998-12-11  
PRIOR APPLICATION NUMBER: 09/291,839  
PRIOR FILING DATE: 1999-04-14  
PRIOR APPLICATION NUMBER: 09/458,457  
PRIOR FILING DATE: 1999-12-10  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 8  
LENGTH: 835  
TYPE: PRT  
ORGANISM: Rattus norvegicus  
US-09-947-199-8

Query Match 16.5%; Score 392.5; DB 10; Length 835;  
Best Local Similarity 29.6%; Pred. No. 4.6e-21;  
Matches 115; Conservative 61; Mismatches 151; Indels 61; Gaps 12;  
QY 11 IKFDDLOFFENCGGSPFSVYRAKVISODKEVAVK-----LTKIEKAEIISLV 59  
DB 458 LQLSIEFHEIIGSSFKYVYKGR--CNKIVAIKRYAANTYCSKSDVDFCREVSLICQ 515  
QY 60 LSHRNIIQFYGVILEPPN-YGIVTEYASLSGLYDINSNSEEMDMHIMTWATDVAKGM 118  
DB 516 LNHPCVVOFVGACLDDBQFAIVTQYISGSLFSLHBRQK-RIIDLGSKLIADVAKGM 574  
QY 119 HYLH-MAEPVKVIHRDLKSRNVVIAADGVKLICDFGASRFRNHTTMSLV--GTFPMA 174

```

Db 575 EYLSLTOP--IIRDLNHNHLLLYEDGHANVADFGESRFLQSLDENMTRKQPNLEWMA 632
QY 175 PEVIT-OSLIVSETCDTYSYGVVLMEMLTREVPFGLEGLOVAMLVKRNRLTIPSSCP 233
Db 633 PEVFOQCRRTYTKADVFSYSLCLWELLTGEIIPFAHLKPAAMAAMVHHIRPPGISIPK 692
QY 234 SFALLHOCWEADAKKRSFKQIISILE-----SMSNDTSLPDKCN--FLHN 279
Db 693 PISLLILGNMAACEGRPEFSEVSKLECCCNELMSPPASNSGSLSPSSSDCLLSR 752
QY 280 KAEMCEIATLEATLERLKLKRLDLSFEKQELKERERLKMWEQKLTQONTPLPLAARMS 339
Db 753 GPGSHVAALRSR-----ELEVALNARSYAGMSQSVGTHSNPGLSLENNRST 802
QY 340 EES-----YFESKTEESN 352
Db 803 QYSTVDKYGVSDPMSLTLLHSRQDDSN 830

```

```

UT 13
09-771-161A-255
; Sequence 255, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 255
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-771-161A-255

```

```

Query Match 16.2%; Score 385; DB 10; Length 1036;
Best Local Similarity 33.1%; Pred. No. 2.1e-20;
Matches 104; Conservative 52; Mismatches 116; Indels 42; Gaps 13;

```

```

Db 15 DLQPFNCGGSGFSGVYRAKW-ISQKEVAVK-KLKIE-----KEAEILSVL 60
QY 125 DLRLLELDGSGFSGVYRGEWDAPSGKTYSVAVKCLKPDVLSQPEAMDFIREVNMHSL 184
QY 61 SHNIIQFYGVILPEPNYGIIVTEYASISGLYDYINSRSEMDMHIMTAWDVAKGMHY 120
Db 185 DHRNLRLHYGVLTLP-MKRVTELAPLGSLDLRLKHQGHFL-LGTLRYAVQVABGMGY 242
QY 121 LHMEAPVKVIRDLKSNVVIADGVKICDPGASR-FHNHTHMLVG-----TFPMAP 175
Db 243 LESK---RFIHRDLAARNLLATRDIVKIGDFGLMRALPONDHYVVOEHRKVPFAMCAP 299
QY 176 EVIOSLVSFSCDTYSYGVVLMEMLT-REVPFGLEGLOVAMLVKRNRLTIPSSCP 234
Db 300 ESKLRTFSHSDTWMFGVLTMEFTYGGEPWIGLNSQILHKKIDKRGELRPPEDCPQD 359
QY 235 FAELHOCWEADAKKRSFKQIIS-IIESMSNDTSL-----PDKCN-----SFLANK 280
Db 360 IYNNWVQCMARKPEDRPTFVALRDFLEAOPTDWRALQDFEEDPKLHIQNDVITYIEGR 419
QY 281 AE---WRCEIATL 291
Db 420 AENYWMRGONTRTL 433

```

```

RESULT 14
US-09-771-161A-256
; Sequence 256, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 256
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-771-161A-256

```

```

Query Match 16.2%; Score 385; DB 10; Length 1036;
Best Local Similarity 33.1%; Pred. No. 2.1e-20;
Matches 104; Conservative 52; Mismatches 116; Indels 42; Gaps 13;

```

```

QY 15 DLQPFNCGGSGFSGVYRAKW-ISQKEVAVK-KLKIE-----KEAEILSVL 60
Db 125 DLRLLELDGSGFSGVYRGEWDAPSGKTYSVAVKCLKPDVLSQPEAMDFIREVNMHSL 184
QY 61 SHNIIQFYGVILPEPNYGIIVTEYASISGLYDYINSRSEMDMHIMTAWDVAKGMHY 120
Db 185 DHRNLRLHYGVLTLP-MKRVTELAPLGSLDLRLKHQGHFL-LGTLRYAVQVABGMGY 242
QY 121 LHMEAPVKVIRDLKSNVVIADGVKICDPGASR-FHNHTHMLVG-----TFPMAP 175
Db 243 LESK---RFIHRDLAARNLLATRDIVKIGDFGLMRALPONDHYVVOEHRKVPFAMCAP 299
QY 176 EVIOSLVSFSCDTYSYGVVLMEMLT-REVPFGLEGLOVAMLVKRNRLTIPSSCP 234
Db 300 ESKLRTFSHSDTWMFGVLTMEFTYGGEPWIGLNSQILHKKIDKRGELRPPEDCPQD 359
QY 235 FAELHOCWEADAKKRSFKQIIS-IIESMSNDTSL-----PDKCN-----SFLANK 280
Db 360 IYNNWVQCMARKPEDRPTFVALRDFLEAOPTDWRALQDFEEDPKLHIQNDVITYIEGR 419
QY 281 AE---WRCEIATL 291
Db 420 AENYWMRGONTRTL 433

```

```

RESULT 15
US-08-987-689A-32
; Sequence 32, Application US/08987689A
; Patent No. US20020048782A1
; GENERAL INFORMATION:
; APPLICANT: Sima Lev
; APPLICANT: Joseph Schlessinger
; TITLE OF INVENTION: PYK2 RELATED PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESS: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0

```



SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/987,689A  
FILING DATE: December 9, 1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/032,824  
FILING DATE: December 11, 1996  
APPLICATION NUMBER: 08/460,626  
FILING DATE: June 2, 1995  
APPLICATION NUMBER: 08/357,642  
FILING DATE: December 15, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 230/110  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-987-689A-32

Query Match 16.1%; Score 382.5; DB 8; Length 251;  
Best Local Similarity 33.9%; Pred. No. 5,8e-21;  
Matches 83; Conservative 51; Mismatches 96; Indels 15; Gaps 7;  
QY 23 GGSFSGSVYRAKWIISQDKVAVKKL-----LKIE---KEAFILSVLSHRNIIQFYGVILEP 75  
DB 7 GGGQYGEVYEGWKAKSLVAVAKTLKEDTVEEFLKEAAVMKEIKHPNLVOLLGCTRE 66  
QY 76 PNYGIETEVASLGLVDYINSNRSEMDMDHIMTATDVAKGMHYLIHMEAPVKVIHRDLK 135  
DB 67 PPFYITTEFTYGNLLDYLRECNROEVNAVVLVMAIQISSAMEYLEKK--NFIHRDLA 123  
QY 136 SRNVYIADGVKICDFGASRPHNHTHMSLVGT--P--MMAPEVIOQLPVSETCDTYSY 192  
DB 124 ANNCVGENHVLKVDVDFGLSLMTGDTYTAHGAKEPIKMTAPESLAVNKFISIKSDWAF 183  
QY 193 GWLWEMLTREV--PFKGLEGLQVAVLVVEKNERLTPSSCPSPFAELHQCWEADAKRP 251  
DB 184 GVLWEIATYGMSPYGIIDRSQV-YELLEKDYRMKRPFGCEPEKVEYELMRACWQNPSPDRP 242  
252 SPKQI 256  
DB 243 SFAEI 247

Search completed: May 1, 2003, 20:52:12  
Job time : 35.1911 secs



GenCore version 5.1.4 p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:33:12 ; Search time 30.4603 Seconds

(without alignment)  
439.505 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 2375  
Sequence: 1 MSSIGASFVQIKFDLQFFD...GDDDDDDDEEDNDMDNSR 455

Scoring table:

BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 2942292 residues

al number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/6D.COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2375	100.0	455	3	US-09-221-235-5
2	2375	100.0	455	3	US-09-221-235-5
3	2375	100.0	455	4	US-09-221-235-5
4	2375	100.0	455	4	US-09-221-235-5
5	2375	100.0	455	4	US-09-221-235-5
6	2375	100.0	455	4	US-09-221-235-5
7	2375	100.0	455	4	US-09-221-235-5
8	2375	100.0	455	4	US-09-221-235-5
9	2375	100.0	455	4	US-09-221-235-5
10	2375	100.0	455	4	US-09-221-235-5
11	2375	100.0	455	4	US-09-221-235-5
12	2375	100.0	455	4	US-09-221-235-5
13	2375	100.0	455	4	US-09-221-235-5
14	2375	100.0	455	4	US-09-221-235-5
15	2375	100.0	455	4	US-09-221-235-5
16	2375	100.0	455	4	US-09-221-235-5
17	2375	100.0	455	4	US-09-221-235-5
18	2375	100.0	455	4	US-09-221-235-5
19	2375	100.0	455	4	US-09-221-235-5
20	2375	100.0	455	4	US-09-221-235-5
21	2375	100.0	455	4	US-09-221-235-5
22	2375	100.0	455	4	US-09-221-235-5
23	2375	100.0	455	4	US-09-221-235-5
24	2375	100.0	455	4	US-09-221-235-5
25	2375	100.0	455	4	US-09-221-235-5
26	2375	100.0	455	4	US-09-221-235-5
27	2375	100.0	455	4	US-09-221-235-5

28	379.5	16.0	505	1	US-08-222-616-20	Sequence 20, Appl
29	379.5	16.0	505	4	US-08-446-648-20	Sequence 20, Appl
30	379.5	16.0	505	4	PCT-US95-04228-20	Sequence 20, Appl
31	377.5	15.9	506	4	US-08-426-509A-6	Sequence 6, Appl
32	377.5	15.9	511	5	PCT-US95-05008-6	Sequence 6, Appl
33	377.5	15.9	511	5	US-09-291-839-2	Sequence 2, Appl
34	376.5	15.9	537	4	US-08-426-509A-11	Sequence 11, Appl
35	376.5	15.9	537	4	PCT-US95-05008-11	Sequence 11, Appl
36	376.5	15.8	283	2	US-08-701-191A-41	Sequence 41, Appl
37	374.5	15.8	259	2	US-07-857-224B-52	Sequence 52, Appl
38	365.5	15.4	536	4	US-08-426-509A-12	Sequence 12, Appl
39	365.5	15.4	536	5	PCT-US95-05008-12	Sequence 12, Appl
40	364	15.3	820	1	US-08-166-717D-6	Sequence 6, Appl
41	358.5	15.1	675	4	US-08-426-509A-4	Sequence 4, Appl
42	358.5	15.1	675	5	PCT-US95-05008-4	Sequence 4, Appl
43	358	15.1	729	1	US-07-640-029-3	Sequence 3, Appl
44	358	15.1	731	1	US-07-921-807B-5	Sequence 5, Appl
45	358	15.1	731	1	US-08-441-944A-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1	US-09-221-235-5	Sequence 5, Application US/09221235	Patent No. 6043040	GENERAL INFORMATION:					
APPLICANT:	Acton, Susan	TITLE OF INVENTION:	NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR	FILE REFERENCE:	NMI-050				
CURRENT FILING DATE:	1998-12-28	EARLIER APPLICATION NUMBER:	09/163,115	NUMBER OF SEQ ID NOS:	15				
SOFTWARE:	Patentin Ver. 2.0	SEQ ID NO 5	LENGTH: 455	TYPE:	PRT				
ORGANISM:	Homo sapiens	US-09-221-235-5	Query Match	100.0%; Score 2375; DB 3; Length 455;	Best Local Similarity	100.0%; Pred. No. 2.3e-192;			
Matches	455;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	MSSIGASFVQIKFDLQFFD	CGGSGFSGSYRAKMSQDXEAVKLLKTEKEAIIISVL	60	DB	1	MSSIGASFVQIKFDLQFFD	CGGSGFSGSYRAKMSQDXEAVKLLKTEKEAIIISVL	60
QY	61	SHRNIOFYGVIIPEPVYGVITEYASISGLSYDYINSRSEMDHMTATVAKGMHY	120	DB	61	SHRNIOFYGVIIPEPVYGVITEYASISGLSYDYINSRSEMDHMTATVAKGMHY	120		
QY	121	LHMEAPYKVIHRDLKSNRVVIADGVLCIDFGASRPHNTHMSLVGTTPMNAPEVIQS	180	DB	121	LHMEAPYKVIHRDLKSNRVVIADGVLCIDFGASRPHNTHMSLVGTTPMNAPEVIQS	180		
QY	181	LPVSETCDITYSYGVVIMEMITREVPFKGLGLOVAMLVKKNRLLTPSSGCPSPFAELH	240	DB	181	LPVSETCDITYSYGVVIMEMITREVPFKGLGLOVAMLVKKNRLLTPSSGCPSPFAELH	240		
QY	241	QCWEADAKKPSFKQIISILSMSNDTSLPDKNSFLHNKAEWRCEIATLEKLEKLERD	300	DB	241	QCWEADAKKPSFKQIISILSMSNDTSLPDKNSFLHNKAEWRCEIATLEKLEKLERD	300		
QY	301	LSTKEQELKERERRLKMEQKLTQSGNTPLLPPLAARMSESYFESKTEBSNAEMSCQI	360	DB	301	LSTKEQELKERERRLKMEQKLTQSGNTPLLPPLAARMSESYFESKTEBSNAEMSCQI	360		
QY	361	TATSNBEGHGNBPLQAMLMGFGDITSNMKAGAVVHSGQIMQAKONSSTKTSKRGGK	420	DB	361	TATSNBEGHGNBPLQAMLMGFGDITSNMKAGAVVHSGQIMQAKONSSTKTSKRGGK	420		

Db 361 TATNGEGHGNPDLQAMLMGFDIFSMNKAAGVMSGMQINNOAKONSCKTTSKRGK 420  
QY 421 KVNMLGSPDFDLSBGDDDDDDGDEEDNDMDNSE 455  
Db 421 KVNMLGSPDFDLSBGDDDDDDGDEEDNDMDNSE 455

## RESULT 2

US-09-221-928-5  
; Sequence 5, Application US/09221928

; Patent No. 6121030  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: NMI-050  
; CURRENT APPLICATION NUMBER: US/09/221,928  
; EARLIER FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 5  
; LENGTH: 455

; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-928-5

Query Match 100.0%; Score 2375; DB 3; Length 455;

Best Local Similarity 100.0%; Pred. No. 2,3e-192;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSIGASFVQIKFDDLOFPENCGGSGFSYVRAKWIISQDEKAVAVKLLIKIEKAEILISVL 60  
Db 1 MSSIGASFVQIKFDDLOFPENCGGSGFSYVRAKWIISQDEKAVAVKLLIKIEKAEILISVL 60  
QY 61 SHRNIIOFYGVILPEPPYGI VTEYASLSGLYDIYINSRSEEMDMHIMTWATDVAKGMY 120  
Db 61 SHRNIIOFYGVILPEPPYGI VTEYASLSGLYDIYINSRSEEMDMHIMTWATDVAKGMY 120  
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTMSLVGTFPMMAPEVIQS 180  
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTMSLVGTFPMMAPEVIQS 180  
QY 181 LPVSETCDITYSGVYLMEMLTREVPFGLEGLOVAMLVKNERLTPSSCPSPFAELH 240  
Db 181 LPVSETCDITYSGVYLMEMLTREVPFGLEGLOVAMLVKNERLTPSSCPSPFAELH 240  
QY 241 QCMWADAKKRPSPFOIISLESMSNDTSLPDKNSFLHNKAERCEIATLERLKLKLERD 300  
Db 241 QCMWADAKKRPSPFOIISLESMSNDTSLPDKNSFLHNKAERCEIATLERLKLKLERD 300  
QY 301 LSPKEOLKERERRLKMEQKLTQOSNTPLLLPLAARMSSESYFESTESNSAEMSCOI 360  
Db 301 LSPKEOLKERERRLKMEQKLTQOSNTPLLLPLAARMSSESYFESTESNSAEMSCOI 360  
QY 361 TATNGEGHGNPDLQAMLMGFDIFSMNKAAGVMSGMQINNOAKONSCKTTSKRGK 420  
Db 361 TATNGEGHGNPDLQAMLMGFDIFSMNKAAGVMSGMQINNOAKONSCKTTSKRGK 420  
QY 421 KVNMLGSPDFDLSBGDDDDDDGDEEDNDMDNSE 455  
Db 421 KVNMLGSPDFDLSBGDDDDDDGDEEDNDMDNSE 455

## RESULT 3

US-09-221-527-5  
; Sequence 5, Application US/09221527

; Patent No. 6146832  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: NMI-050  
; CURRENT APPLICATION NUMBER: US/09/221,527

; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-527-5

## Query Match

100.0%; Score 2375; DB 4; Length 455;

Best Local Similarity 100.0%; Pred. No. 2,3e-192;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSIGASFVQIKFDDLOFPENCGGSGFSYVRAKWIISQDEKAVAVKLLIKIEKAEILISVL 60  
Db 1 MSSIGASFVQIKFDDLOFPENCGGSGFSYVRAKWIISQDEKAVAVKLLIKIEKAEILISVL 60  
QY 61 SHRNIIOFYGVILPEPPYGI VTEYASLSGLYDIYINSRSEEMDMHIMTWATDVAKGMY 120  
Db 61 SHRNIIOFYGVILPEPPYGI VTEYASLSGLYDIYINSRSEEMDMHIMTWATDVAKGMY 120  
QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTMSLVGTFPMMAPEVIQS 180  
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTMSLVGTFPMMAPEVIQS 180  
QY 181 LPVSETCDITYSGVYLMEMLTREVPFGLEGLOVAMLVKNERLTPSSCPSPFAELH 240  
Db 181 LPVSETCDITYSGVYLMEMLTREVPFGLEGLOVAMLVKNERLTPSSCPSPFAELH 240  
QY 241 QCMWADAKKRPSPFOIISLESMSNDTSLPDKNSFLHNKAERCEIATLERLKLKLERD 300  
Db 241 QCMWADAKKRPSPFOIISLESMSNDTSLPDKNSFLHNKAERCEIATLERLKLKLERD 300  
QY 301 LSPKEOLKERERRLKMEQKLTQOSNTPLLLPLAARMSSESYFESTESNSAEMSCOI 360  
Db 301 LSPKEOLKERERRLKMEQKLTQOSNTPLLLPLAARMSSESYFESTESNSAEMSCOI 360  
QY 361 TATNGEGHGNPDLQAMLMGFDIFSMNKAAGVMSGMQINNOAKONSCKTTSKRGK 420  
Db 361 TATNGEGHGNPDLQAMLMGFDIFSMNKAAGVMSGMQINNOAKONSCKTTSKRGK 420  
QY 421 KVNMLGSPDFDLSBGDDDDDDGDEEDNDMDNSE 455  
Db 421 KVNMLGSPDFDLSBGDDDDDDGDEEDNDMDNSE 455

## RESULT 4

US-09-221-236-5  
; Sequence 5, Application US/09221236

; Patent No. 6146841  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: NMI-050  
; CURRENT APPLICATION NUMBER: US/09/221,236  
; EARLIER FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 5  
; LENGTH: 455

; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-236-5

Query Match 100.0%; Score 2375; DB 4; Length 455;

Best Local Similarity 100.0%; Pred. No. 2,3e-192;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSIGASFVQIKFDDLOFPENCGGSGFSYVRAKWIISQDEKAVAVKLLIKIEKAEILISVL 60

Db 1 MSSLGASFVQIKFPDLOFFENCGGSGFSYVRARAKISQDKEVAVKLLIKEAEIISVL 60  
Qy 61 SHRNIIQFYGVILEPPNYGIVTEYASLGSLYDYINSRSEEMDMHIMTATDVAKGMY 120  
Db 61 SHRNIIQFYGVILEPPNYGIVTEYASLGSLYDYINSRSEEMDMHIMTATDVAKGMY 120  
Qy 121 LHMEAPVKVIRHDLKSRNVVIAADGVLCIDFGASRFPHNHTTHMSLVGTFFPMAPEVIQS 180  
Db 121 LHMEAPVKVIRHDLKSRNVVIAADGVLCIDFGASRFPHNHTTHMSLVGTFFPMAPEVIQS 180  
Qy 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAMLVENKNERLTTPSSCPRSFAELIH 240  
Db 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAMLVENKNERLTTPSSCPRSFAELIH 240  
Qy 241 QCWEADAKKRPSPFOIISILESMSNDTSLPDKCNSFLHNKAEMRCEIATLERLKLKLERD 300  
Db 241 QCWEADAKKRPSPFOIISILESMSNDTSLPDKCNSFLHNKAEMRCEIATLERLKLKLERD 300  
Qy 301 LSPFEOELKERERRLKWEQKLTQOSNTPLLLPLAARMSSESYFESKTEBSNSAEMSCQI 360  
Db 301 LSPFEOELKERERRLKWEQKLTQOSNTPLLLPLAARMSSESYFESKTEBSNSAEMSCQI 360  
Qy 361 TATNSGCHGNPSSLQAMMLMGFGDIFSMNKAAGAVMHSQMIOINQAKONSSTKTSKRGK 420  
Db 361 TATNSGCHGNPSSLQAMMLMGFGDIFSMNKAAGAVMHSQMIOINQAKONSSTKTSKRGK 420  
Qy 421 KVNMLGFSDFDLSGDDDDDDGEEEDNDMDNSE 455  
Db 421 KVNMLGFSDFDLSGDDDDDDGEEEDNDMDNSE 455

RESULT 5  
US-09-221-416-5  
; Sequence 5, Application US/09221416  
; Patent No. 6153417  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: NMI-050  
; CURRENT APPLICATION NUMBER: US/09/221,416  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 5  
LENGTH: 455  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-221-416-5

Query Match 100.0%; Score 2375; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 2,3e-192;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MSSLGASFVQIKFPDLOFFENCGGSGFSYVRARAKISQDKEVAVKLLIKEAEIISVL 60  
Db 1 MSSLGASFVQIKFPDLOFFENCGGSGFSYVRARAKISQDKEVAVKLLIKEAEIISVL 60  
Qy 61 SHRNIIQFYGVILEPPNYGIVTEYASLGSLYDYINSRSEEMDMHIMTATDVAKGMY 120  
Db 61 SHRNIIQFYGVILEPPNYGIVTEYASLGSLYDYINSRSEEMDMHIMTATDVAKGMY 120  
Qy 121 LHMEAPVKVIRHDLKSRNVVIAADGVLCIDFGASRFPHNHTTHMSLVGTFFPMAPEVIQS 180  
Db 121 LHMEAPVKVIRHDLKSRNVVIAADGVLCIDFGASRFPHNHTTHMSLVGTFFPMAPEVIQS 180  
Qy 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAMLVENKNERLTTPSSCPRSFAELIH 240  
Db 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAMLVENKNERLTTPSSCPRSFAELIH 240  
Qy 241 QCWEADAKKRPSPFOIISILESMSNDTSLPDKCNSFLHNKAEMRCEIATLERLKLKLERD 300  
Db 241 QCWEADAKKRPSPFOIISILESMSNDTSLPDKCNSFLHNKAEMRCEIATLERLKLKLERD 300

Db 241 QCWEADAKKRPSPFOIISILESMSNDTSLPDKCNSFLHNKAEMRCEIATLERLKLKLERD 300  
Qy 301 LSPFEOELKERERRLKWEQKLTQOSNTPLLLPLAARMSSESYFESKTEBSNSAEMSCQI 360  
Db 301 LSPFEOELKERERRLKWEQKLTQOSNTPLLLPLAARMSSESYFESKTEBSNSAEMSCQI 360  
Qy 361 TATNSGCHGNPSSLQAMMLMGFGDIFSMNKAAGAVMHSQMIOINQAKONSSTKTSKRGK 420  
Db 361 TATNSGCHGNPSSLQAMMLMGFGDIFSMNKAAGAVMHSQMIOINQAKONSSTKTSKRGK 420  
Qy 421 KVNMLGFSDFDLSGDDDDDDGEEEDNDMDNSE 455  
Db 421 KVNMLGFSDFDLSGDDDDDDGEEEDNDMDNSE 455

RESULT 6  
US-09-221-245-5  
; Sequence 5, Application US/09221245  
; Patent No. 6180358  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: NMI-050  
; CURRENT APPLICATION NUMBER: US/09/221,245  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: US 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 5  
LENGTH: 455  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-221-245-5

Query Match 100.0%; Score 2375; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 2,3e-192;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MSSLGASFVQIKFPDLOFFENCGGSGFSYVRARAKISQDKEVAVKLLIKEAEIISVL 60  
Db 1 MSSLGASFVQIKFPDLOFFENCGGSGFSYVRARAKISQDKEVAVKLLIKEAEIISVL 60  
Qy 61 SHRNIIQFYGVILEPPNYGIVTEYASLGSLYDYINSRSEEMDMHIMTATDVAKGMY 120  
Db 61 SHRNIIQFYGVILEPPNYGIVTEYASLGSLYDYINSRSEEMDMHIMTATDVAKGMY 120  
Qy 121 LHMEAPVKVIRHDLKSRNVVIAADGVLCIDFGASRFPHNHTTHMSLVGTFFPMAPEVIQS 180  
Db 121 LHMEAPVKVIRHDLKSRNVVIAADGVLCIDFGASRFPHNHTTHMSLVGTFFPMAPEVIQS 180  
Qy 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAMLVENKNERLTTPSSCPRSFAELIH 240  
Db 181 LPVSETCDTYSYGVVLEWMLTREVPFKGLEGLQVAMLVENKNERLTTPSSCPRSFAELIH 240  
Qy 241 QCWEADAKKRPSPFOIISILESMSNDTSLPDKCNSFLHNKAEMRCEIATLERLKLKLERD 300  
Db 241 QCWEADAKKRPSPFOIISILESMSNDTSLPDKCNSFLHNKAEMRCEIATLERLKLKLERD 300  
Qy 301 LSPFEOELKERERRLKWEQKLTQOSNTPLLLPLAARMSSESYFESKTEBSNSAEMSCQI 360  
Db 301 LSPFEOELKERERRLKWEQKLTQOSNTPLLLPLAARMSSESYFESKTEBSNSAEMSCQI 360  
Qy 361 TATNSGCHGNPSSLQAMMLMGFGDIFSMNKAAGAVMHSQMIOINQAKONSSTKTSKRGK 420  
Db 361 TATNSGCHGNPSSLQAMMLMGFGDIFSMNKAAGAVMHSQMIOINQAKONSSTKTSKRGK 420  
Qy 421 KVNMLGFSDFDLSGDDDDDDGEEEDNDMDNSE 455  
Db 421 KVNMLGFSDFDLSGDDDDDDGEEEDNDMDNSE 455

```
RESULT 7
US-09-163-115-5
; Sequence 5, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/163.115A
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-163-115-5

Query Match          100.0%; Score 2375; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.3e-192;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSIGASFGVQIKFPDLOFFENCGGSGFSYVRRAKWIISQDKKAVAVKLLKIEKAEIISVL 60
DB 1 MSSIGASFGVQIKFPDLOFFENCGGSGFSYVRRAKWIISQDKKAVAVKLLKIEKAEIISVL 60
QY 61 SHRNIIQFYGVILEPPNYGIVTEYASIGSLYDYNINSRSEEMDMHIMTATDVAKGMY 120
DB 61 SHRNIIQFYGVILEPPNYGIVTEYASIGSLYDYNINSRSEEMDMHIMTATDVAKGMY 120
QY 121 LHMEAPVKVVIHRDLKSRNVVIAADGVLCICDYGASRFHNHTTHSLVGTPEPMAPEVIQS 180
DB 121 LHMEAPVKVVIHRDLKSRNVVIAADGVLCICDYGASRFHNHTTHSLVGTPEPMAPEVIQS 180
QY 181 LPVSETCDTYSYGVILWEMLTREVPFKGLEQVAMLVKNERLTIPSSCPSPFAELH 240
DB 181 LPVSETCDTYSYGVILWEMLTREVPFKGLEQVAMLVKNERLTIPSSCPSPFAELH 240
QY 241 QCEWADAKKRPSPFOIISILESMNDTSLPDKNSFLHNKAEMRCIEATLERLKLKLERD 300
DB 241 QCEWADAKKRPSPFOIISILESMNDTSLPDKNSFLHNKAEMRCIEATLERLKLKLERD 300
QY 301 LSPFEOELKERRERLKWEQKLTQOSTPLLLPLAARMSSESYFESKTESNSAEMSCQI 360
DB 301 LSPFEOELKERRERLKWEQKLTQOSTPLLLPLAARMSSESYFESKTESNSAEMSCQI 360
QY 361 TATNGEGHGNPISLOAMLMGFGDIFSNKAGAVMHSQOINMOAKONSSKTTSKRGK 420
DB 361 TATNGEGHGNPISLOAMLMGFGDIFSNKAGAVMHSQOINMOAKONSSKTTSKRGK 420
QY 421 KVNNALGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455
DB 421 KVNNALGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455

RESULT 8
US-09-221-528-5
; Sequence 5, Application US/09221528
; Patent No. 6190874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221.528
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163.115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
US-09-221-528-5

Query Match          100.0%; Score 2375; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.3e-192;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSIGASFGVQIKFPDLOFFENCGGSGFSYVRRAKWIISQDKKAVAVKLLKIEKAEIISVL 60
DB 1 MSSIGASFGVQIKFPDLOFFENCGGSGFSYVRRAKWIISQDKKAVAVKLLKIEKAEIISVL 60
QY 61 SHRNIIQFYGVILEPPNYGIVTEYASIGSLYDYNINSRSEEMDMHIMTATDVAKGMY 120
DB 61 SHRNIIQFYGVILEPPNYGIVTEYASIGSLYDYNINSRSEEMDMHIMTATDVAKGMY 120
QY 121 LHMEAPVKVVIHRDLKSRNVVIAADGVLCICDYGASRFHNHTTHSLVGTPEPMAPEVIQS 180
DB 121 LHMEAPVKVVIHRDLKSRNVVIAADGVLCICDYGASRFHNHTTHSLVGTPEPMAPEVIQS 180
QY 181 LPVSETCDTYSYGVILWEMLTREVPFKGLEQVAMLVKNERLTIPSSCPSPFAELH 240
DB 181 LPVSETCDTYSYGVILWEMLTREVPFKGLEQVAMLVKNERLTIPSSCPSPFAELH 240
QY 241 QCEWADAKKRPSPFOIISILESMNDTSLPDKNSFLHNKAEMRCIEATLERLKLKLERD 300
DB 241 QCEWADAKKRPSPFOIISILESMNDTSLPDKNSFLHNKAEMRCIEATLERLKLKLERD 300
QY 301 LSPFEOELKERRERLKWEQKLTQOSTPLLLPLAARMSSESYFESKTESNSAEMSCQI 360
DB 301 LSPFEOELKERRERLKWEQKLTQOSTPLLLPLAARMSSESYFESKTESNSAEMSCQI 360
QY 361 TATNGEGHGNPISLOAMLMGFGDIFSNKAGAVMHSQOINMOAKONSSKTTSKRGK 420
DB 361 TATNGEGHGNPISLOAMLMGFGDIFSNKAGAVMHSQOINMOAKONSSKTTSKRGK 420
QY 421 KVNNALGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455
DB 421 KVNNALGFSDFDLSEGGDDDDDDGEEEDNDMDNSE 455

RESULT 9
US-09-593-553-5
; Sequence 5, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/593.553
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/163.115
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-593-553-5

Query Match          100.0%; Score 2375; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 2.3e-192;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSIGASFGVQIKFPDLOFFENCGGSGFSYVRRAKWIISQDKKAVAVKLLKIEKAEIISVL 60
DB 1 MSSIGASFGVQIKFPDLOFFENCGGSGFSYVRRAKWIISQDKKAVAVKLLKIEKAEIISVL 60
QY 61 SHRNIIQFYGVILEPPNYGIVTEYASIGSLYDYNINSRSEEMDMHIMTATDVAKGMY 120
DB 61 SHRNIIQFYGVILEPPNYGIVTEYASIGSLYDYNINSRSEEMDMHIMTATDVAKGMY 120
QY 121 LHMEAPVKVVIHRDLKSRNVVIAADGVLCICDYGASRFHNHTTHSLVGTPEPMAPEVIQS 180
DB 121 LHMEAPVKVVIHRDLKSRNVVIAADGVLCICDYGASRFHNHTTHSLVGTPEPMAPEVIQS 180
```

Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHSLVGTFFPMAPVYIOS 180  
Qy 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLBGLQVAMLVKNERLTTPSCPRSPAEILH 240  
Db 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLBGLQVAMLVKNERLTTPSCPRSPAEILH 240  
Qy 241 QCWEADAKKRPSPFOIISILESMSNDTSLPDKNSFLHNKAWRCCEIATLERLKLKLERD 300  
Db 241 QCWEADAKKRPSPFOIISILESMSNDTSLPDKNSFLHNKAWRCCEIATLERLKLKLERD 300  
Qy 301 LSPKEOGLKERERLAKMEQKLTQCSNTPLLPLAARMSSESYESTKEESNAEMSCQI 360  
Db 301 LSPKEOGLKERERLAKMEQKLTQCSNTPLLPLAARMSSESYESTKEESNAEMSCQI 360  
Qy 361 TATSNGGCHGNPISLOAMLMGFGDIFSMNKAAGVMSGQINNQAONSKSTTSKRGK 420  
Db 361 TATSNGGCHGNPISLOAMLMGFGDIFSMNKAAGVMSGQINNQAONSKSTTSKRGK 420  
Qy 421 KVNNALGFSDFDSEGDGDDDDDEEBEDNDMNSE 455  
Db 421 KVNNALGFSDFDSEGDGDDDDDEEBEDNDMNSE 455

## RESULT 10

US-09-221-237-5  
Sequence 5, Application US/09221237  
Patent No. 6214597  
GENERAL INFORMATION:  
APPLICANT: Acton, Susan  
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
FILE REFERENCE: MRI-050  
CURRENT APPLICATION NUMBER: US/09/221,237  
EARLIER FILING DATE: 1998-12-28  
EARLIER APPLICATION NUMBER: 09/163,115  
NUMBER OR SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 455  
TYPE: PR1  
ORGANISM: Homo sapiens  
US-09-221-237-5

Query Match 100.0%; Score 2375; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 2.3e-192;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MSSLGASFPVQIKFPDLOFFENCGGSGFSVYRAKMIQDKEVAVKLLIKEAEILSVL 60  
1 MSSLGASFPVQIKFPDLOFFENCGGSGFSVYRAKMIQDKEVAVKLLIKEAEILSVL 60  
Qy 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDIYNSNRSEMDMDHMTATVAKGMHY 120  
Db 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDIYNSNRSEMDMDHMTATVAKGMHY 120  
Qy 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHSLVGTFFPMAPVYIOS 180  
Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHSLVGTFFPMAPVYIOS 180  
Qy 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLBGLQVAMLVKNERLTTPSCPRSPAEILH 240  
Db 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLBGLQVAMLVKNERLTTPSCPRSPAEILH 240  
Qy 241 QCWEADAKKRPSPFOIISILESMSNDTSLPDKNSFLHNKAWRCCEIATLERLKLKLERD 300  
Db 241 QCWEADAKKRPSPFOIISILESMSNDTSLPDKNSFLHNKAWRCCEIATLERLKLKLERD 300  
Qy 301 LSPKEOGLKERERLAKMEQKLTQCSNTPLLPLAARMSSESYESTKEESNAEMSCQI 360  
Db 301 LSPKEOGLKERERLAKMEQKLTQCSNTPLLPLAARMSSESYESTKEESNAEMSCQI 360  
Qy 361 TATSNGGCHGNPISLOAMLMGFGDIFSMNKAAGVMSGQINNQAONSKSTTSKRGK 420  
Db 361 TATSNGGCHGNPISLOAMLMGFGDIFSMNKAAGVMSGQINNQAONSKSTTSKRGK 420

Db 361 TATSNGGCHGNPISLOAMLMGFGDIFSMNKAAGVMSGQINNQAONSKSTTSKRGK 420  
Qy 421 KVNNALGFSDFDSEGDGDDDDDEEBEDNDMNSE 455  
Db 421 KVNNALGFSDFDSEGDGDDDDDEEBEDNDMNSE 455

## RESULT 11

US-08-205-018-2  
Sequence 2, Application US/08205018  
Patent No. 5554523  
GENERAL INFORMATION:  
APPLICANT: Reddy, Usharani R.  
TITLE OF INVENTION: No. 5554523el Protein Kinase, Nucleic Acid  
TITLE OF INVENTION: Sequences Encoding the Same and Methods Related  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and  
ADDRESSEE: No. 5554523is  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/205,018  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Gaumont, Rebecca R.  
REGISTRATION NUMBER: 35,152  
REFERENCE/DOCKET NUMBER: CH-0488  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 668 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-205-018-2

Query Match 23.8%; Score 565.5; DB 1; Length 668;  
Best Local Similarity 37.8%; Pred. No. 1.5e-39;  
Matches 135; Conservative 71; Mismatches 116; Indels 35; Gaps 13;

Qy 10 QIKFDDLOFFENCGGSGFSVYRAKMIQDKEVAVKLLIKEAEI--LSVLSHRNIIQ 67  
Db 119 EVPEEILDLQWQSGAGVFLGRF--HGEEVAVKVRDL-KETDIKHLKLRKHPNIT 175  
Qy 68 FYGVILEPPNYGIVTEYASLSGLYDIYNSNRSEMDMDHMTATVAKGMHYLMEAPV 127  
Db 176 FKVGCTQAPCYCIIIMEFCAGQGLYEVLRAGRPTVPSL--LVDMSMGJAGGNVYHLH-- 230  
Qy 128 KVHRDLKSRNVVIAADGVLCIDFGASR--FHNHTTHSLVGTFFPMAPVYIOSLPVSET 186  
Db 231 KIHRDLKSRNMLITYDVVKISDFGSKLSDKSTGMSFAGTYAMAPVYIRNRPVSEK 290  
Qy 187 CDYISYGVVLMEMLTREVPFKGLBGLQVAMLVKNERLTTPSCPRSPAEILHQCWEAD 246  
Db 291 VDIWISFGVLMELTGBIPYKDVSASAIWGVGNSLHLVPSSCPDGFILRLQCNWSK 350  
Qy 247 AKRPSFKOIIISILESMSNT--SLPDKNSFLHNKAWRCCEIATLERL-----KLE 298  
Db 351 PRNPSFRQILLHLDIASAVLSTPQ--TYFSQAEMREBVKLHFKIKISEGTCLEHLE 408

QY 299 RDLSPKOE-----LKERERLKMMECKLTQSNTPILLPLAARMSSESYESTEE 350  
Db 409 EELVRRERREIRHADRERH---YERKLERANN--LYMEINMLAQ---LELKERE 456

## RESULT 12

US-08-395-580-2

Sequence 2, Application US/08395580

Patent No. 5676945

GENERAL INFORMATION:

APPLICANT: Ushantari R. Reddy, David Pleasure and the Children's

APPLICANT: Hospital of Philadelphia

TITLE OF INVENTION: No. 5676945c1 Protein Kinase, Nucleic Acid

TITLE OF INVENTION: Sequences Encoding the Same and Methods Related Thereto

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5676945c1s

STREET: One Liberty Place - 46th Floor

CITY: Philadelphia

STATE: PA

COUNTRY: U.S.A.

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch disk, 720 Kb

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/395,580

FILING DATE: herewith

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/205,018

FILING DATE: 01-MAR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Rebecca L. Ralph (formerly Gaumond)

REGISTRATION NUMBER: 35,152

REFERENCE/DOCKET NUMBER: CH-0488

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-568-3100

TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 859 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

08-395-580-2

Query Match 23.8%; Score 565.5; DB 1; Length 859;  
Best Local Similarity 37.8%; Pred. No. 2.1e-39;  
Matches 135; Conservative 71; Mismatches 116; Indels 35; Gaps 13;

QY 10 QIKRDDLQFFENCGGSGSVYRAKWIQDKEVAVKLLIKEAEI--LSVLSHRNIQ 67  
Db 119 EVFPEEILDLQWVSGAGVFLGRF--HGEVAVKVRDL-KETDILKRLKLGHPNITT 175  
QY 68 FYGVILPEPNYGIYEVASLGLDYINSNRSEMDMDHMTWATDVAKGMHYLHMEAPV 127  
Db 176 FKVGCTQAPCYCIMEFCAQQLVEVLRAGRPVTPSL--LVDWMSGIAGGNNYHLH--- 230  
QY 128 KVIHRDLKSRNVVIAADGVAKICDYGASR-FHNHTTMSLVGTFFPMAPVYIOSLPVSET 186  
Db 231 KIHRLDKSPMMLITVDVVKISDFTGSKELSDKSTKMSFAGTVAMAPVIRNRPVSEK 290  
QY 187 CDTSYGVVLWEMLTREVPFEGLEGLQVAMLVKNERLTISSCGPRSPFALLHQWEAD 246  
Db 291 VDIWSPGVVLWELLTGEIPYDVDSALIMVGSNSLHLPPVSSCPDGFKILLRQCNMSK 350  
QY 247 AKRPPSPFOIISLESMSNDT-SLPDKCNFLLNKAWEKCEIATELRK-----KLE 298  
Db 351 PRNPPSPFOIILHDIADVLTSTPQ--TYFKSQAEWREBEVKLFEKIKSEGTCLHRL 408

QY 299 RDLSPKOE-----LKERERLKMMECKLTQSNTPILLPLAARMSSESYESTEE 350  
Db 409 EELVRRERREIRHADRERH---YERKLERANN--LYMEINMLAQ---LELKERE 456

## RESULT 13

PCT-US95-02792-2

Sequence 2, Application PC/TUS9502792

GENERAL INFORMATION:

APPLICANT: Ushantari R. Reddy, David Pleasure and the Children's

APPLICANT: Hospital of Philadelphia

TITLE OF INVENTION: Novel Protein Kinase, Nucleic Acid

TITLE OF INVENTION: Sequences Encoding the Same and Methods Related Thereto

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and Norris

STREET: One Liberty Place - 46th Floor

CITY: Philadelphia

STATE: PA

COUNTRY: U.S.A.

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch disk, 720 Kb

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/02792

FILING DATE: herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/205,018

FILING DATE: 01-MAR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Rebecca L. Ralph (formerly Gaumond)

REGISTRATION NUMBER: 35,152

REFERENCE/DOCKET NUMBER: CH-0488

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-568-3100

TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 859 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US95-02792-2

Query Match 23.8%; Score 565.5; DB 5; Length 859;  
Best Local Similarity 37.8%; Pred. No. 2.1e-39;  
Matches 135; Conservative 71; Mismatches 116; Indels 35; Gaps 13;

QY 10 QIKRDDLQFFENCGGSGSVYRAKWIQDKEVAVKLLIKEAEI--LSVLSHRNIQ 67  
Db 119 EVFPEEILDLQWVSGAGVFLGRF--HGEVAVKVRDL-KETDILKRLKLGHPNITT 175  
QY 68 FYGVILPEPNYGIYEVASLGLDYINSNRSEMDMDHMTWATDVAKGMHYLHMEAPV 127  
Db 176 FKVGCTQAPCYCIMEFCAQQLVEVLRAGRPVTPSL--LVDWMSGIAGGNNYHLH--- 230  
QY 128 KVIHRDLKSRNVVIAADGVAKICDYGASR-FHNHTTMSLVGTFFPMAPVYIOSLPVSET 186  
Db 231 KIHRLDKSPMMLITVDVVKISDFTGSKELSDKSTKMSFAGTVAMAPVIRNRPVSEK 290  
QY 187 CDTSYGVVLWEMLTREVPFEGLEGLQVAMLVKNERLTISSCGPRSPFALLHQWEAD 246  
Db 291 VDIWSPGVVLWELLTGEIPYDVDSALIMVGSNSLHLPPVSSCPDGFKILLRQCNMSK 350  
QY 247 AKRPPSPFOIISLESMSNDT-SLPDKCNFLLNKAWEKCEIATELRK-----KLE 298  
Db 351 PRNPPSPFOIILHDIADVLTSTPQ--TYFKSQAEWREBEVKLFEKIKSEGTCLHRL 408



QY 299 RDLSEKOE-----LKERERLKMEOUJTEOSNTPLLLPLAAMSESESYESTER 350  
DB 409 BELVRRREERLHDLIREH-----YERLBERANN--LYMELNMLMQ---LELKERE 456

RESULT 14  
US-07-928-464-2

Sequence 2, Application US/07928464  
Patent No. 5367065  
GENERAL INFORMATION:  
APPLICANT: Ecker, Joseph R.  
APPLICANT: Kieber, Joseph J.  
TITLE OF INVENTION: Constitutive Triple Response Gene and  
TITLE OF INVENTION: Mutations  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz and  
ADDRESSEE: No. 5367065r1s  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/928,464  
FILING DATE: 19920810  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Miller, Suzanne E.  
REGISTRATION NUMBER: 32,279  
REFERENCE/DOCKET NUMBER: UPN-1086  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 821 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-928-464-2

Query Match 20.5%; Score 488; DB 1; Length 821;  
Best Local Similarity 40.7%; Pred. No. 6,9e-33;

Matches 109; Conservative 44; Mismatches 99; Indels 16; Gaps 5;

QY 9 VOIKFDDLOFFENCGGSGFVYRAKWTISQKEVAVKCLKIE-----KEAEITLS 58  
DB 544 MDIPWCDLNIKEKIGAGSGFVHRAEMHGSD--VAVKILMEDQFHAERVNEFLREVAIMK 601  
QY 59 VLSHRNIIQFYGVILPEPPNYGIVTEYASLSLYDIIN--SNSSEEMDMHMTWATDVAKG 117  
DB 602 RLHRPNIVLFMGAVTQPNLSIVTEYLSRGSLYRLHLSGARERQDERRRLSMAYDVAKG 661  
QY 118 MYLHMEAPVKVIHRDLKSRNVVIAADGVLCICPGASRFNNHT--THMSLVGTFPMWAP 175  
DB 662 MYLHNRNP-PIVHRDLKSPNLDVKKTYAVCDGSLRLKASTLSKSAAGTFPMWAP 720  
QY 176 EYIGSLPVSETCDTYSYGVVLMEMLTREVFPKGLGLQVAMLVYKNERLTISSCPRS 235  
DB 721 EYLRDEPENSISDVYSFVILIMELATLQOPWGNLNPQVAAVGFCKRLIIPRNLPQV 780  
QY 236 AELHOCWEADAKKRPSPKQIISLES 263  
DB 781 AAIIEGCWTNEPMKRPSPFATIMDLRL 808

RESULT 15  
US-08-003-311B-2

Sequence 2, Application US/08003311B  
Patent No. 5444166  
GENERAL INFORMATION:  
APPLICANT: Ecker, Joseph R.  
APPLICANT: Kieber, Joseph J.  
TITLE OF INVENTION: Constitutive Triple Response Gene  
TITLE OF INVENTION: and Mutations  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and  
ADDRESSEE: No. 5444166r1s  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/003,311B  
FILING DATE: January 12, 1993  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/928,464  
FILING DATE: August 10, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lori Y. Beardsell  
REGISTRATION NUMBER: 34,293  
REFERENCE/DOCKET NUMBER: UPN-1108  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 821 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-003-311B-2

Query Match 20.5%; Score 488; DB 1; Length 821;  
Best Local Similarity 40.7%; Pred. No. 6,9e-33;  
Matches 109; Conservative 44; Mismatches 99; Indels 16; Gaps 5;

QY 9 VOIKFDDLOFFENCGGSGFVYRAKWTISQKEVAVKCLKIE-----KEAEITLS 58  
DB 544 MDIPWCDLNIKEKIGAGSGFVHRAEMHGSD--VAVKILMEDQFHAERVNEFLREVAIMK 601  
QY 59 VLSHRNIIQFYGVILPEPPNYGIVTEYASLSLYDIIN--SNSSEEMDMHMTWATDVAKG 117  
DB 602 RLHRPNIVLFMGAVTQPNLSIVTEYLSRGSLYRLHLSGARERQDERRRLSMAYDVAKG 661  
QY 118 MYLHMEAPVKVIHRDLKSRNVVIAADGVLCICPGASRFNNHT--THMSLVGTFPMWAP 175  
DB 662 MYLHNRNP-PIVHRDLKSPNLDVKKTYAVCDGSLRLKASTLSKSAAGTFPMWAP 720  
QY 176 EYIGSLPVSETCDTYSYGVVLMEMLTREVFPKGLGLQVAMLVYKNERLTISSCPRS 235  
DB 721 EYLRDEPENSISDVYSFVILIMELATLQOPWGNLNPQVAAVGFCKRLIIPRNLPQV 780  
QY 236 AELHOCWEADAKKRPSPKQIISLES 263  
DB 781 AAIIEGCWTNEPMKRPSPFATIMDLRL 808

Search completed: May 1, 2003, 20:39:29  
Job time: 37.4603 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:31:42 ; Search time 14.8131 Seconds

(without alignments)  
1602.986 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_31\_277

Perfect score: 1300  
Sequence: 1 TRAKWISQDKEVAVKLLIKI.....SILESMSNTSLPDKNSFL 247

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Minimum number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR73.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	ID	Description
1	511.5	39.3	406 2 T52626	probable mitogen-a
2	510.5	39.3	394 2 JU0229	mixed-lineage prot
3	495.5	38.1	328 2 T16747	hypothetical prote
4	490.5	37.7	954 1 S68178	mixed-lineage prote
5	490	37.7	668 2 JC2363	protein kinase (EC
6	490	37.7	888 2 A55318	serine/threonine p
7	488	37.5	888 2 JC5399	dual leucine zippe
8	482.5	37.1	847 1 A53800	MAPK delta-1 prot
9	480	36.9	886 2 T48544	mixed-lineage prot
10	475.5	36.6	1030 2 S29851	hypothetical prote
11	460.5	35.4	462 2 S29851	protein kinase 6 (
12	457.5	35.2	390 2 T01451	protein kinase hom
13	453	34.8	982 2 T06576	probable protein k
14	448	34.5	412 2 T10671	protein kinase hom
15	442	34.0	821 2 T48400	serine/threonine-p
16	436.5	33.6	1015 2 T00726	probable serine/th
17	434	33.4	848 2 B87950	protein F3352.2 [1
18	434	33.4	855 2 T20082	hypothetical prote
19	431.5	33.2	553 2 T04683	hypothetical prote
20	427	33.2	829 2 T07406	probable protein k
21	427	32.8	475 2 T12955	probable protein k
22	420.5	32.3	546 2 D84555	probable protein k
23	418	32.2	407 2 G84635	probable protein k
24	417	32.1	736 2 T05137	protein kinase hom
25	403.5	31.0	988 2 F86316	protein T10022.13
26	402	30.9	963 2 T09911	probable serine/th
27	396.5	30.5	545 2 T05675	hypothetical prote
28	395	30.4	567 2 JC5957	transforming growt
29	395	30.4	579 2 JC5955	transforming growt

30	395	30.4	606 2 JC5956	transforming growt
31	381.5	29.3	738 2 P96701	hypothetical prote
32	369.5	28.4	1130 1 T00486	protein-tyrosine k
33	369	28.4	1257 2 T00486	serine/threonine-s
34	368.5	28.3	981 1 F0MVGW	gag-abl polyprotei
35	368.5	28.3	1123 2 A39962	kinase-related tra
36	367	28.2	364 2 G71410	probable protein k
37	365.5	28.1	1146 2 B35962	protein-tyrosine k
38	365.5	28.1	1182 2 A35962	protein-tyrosine k
39	364.5	28.0	1520 1 TVFPA	protein-tyrosine k
40	362.5	27.9	357 2 C84856	probable protein k
41	362.5	27.9	391 2 T48115	protein kinase ATM
42	362	27.8	842 2 T32258	hypothetical prote
43	358.5	27.6	1147 2 F86297	hypothetical prote
44	355.5	27.3	1584 2 T18276	protein-tyrosine k
45	354.5	27.3	1171 2 T12956	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

T52626

probable mitogen-activated protein kinase MAPK delta-1 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #ext\_change 08-Dec-2000

C/Accession: T52626

R/Jouanic, S.; Hamal, A.; Leprince, A.S.; Tregear, J.W.; Krets, M.; Henry, Y.

Gene 229, 171-81, 1999

A/Title: Characterization of novel plant genes encoding MEKK/STK1 and RAF-related prot:

A/Reference number: Z24447; MUID:9916996; PMID:10095117

A/Accession: T52626

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-406 <100>

A/Cross-references: EMBL:Y14199; NID:G2253009; PIDN:CA474591.1; PID:G2253010

A/Experimental source: cultivar Columbia

C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol.

Query Match 39.3%; Score 511.5; DB 2; Length 406;

Best Local Similarity 40.5%; Pred. No. 1.8e-23;

Matches 104; Conservative 51; Mismatches 85; Indels 17; Gaps 6;

QY	1	YRAKWSIQDKEVAVKLL-----KIKKAEILSVSHNIIQFVILEPPNYGI 50
DB	150	YRAEW--NGTEVAVKFLDDPFGDALTQPKSEITMLRIKHNVVLFMGAVTRPPNFSI 207
QY	51	VTEVASIGSLVDYINSRSEMDIMTWATDVAKGMHYLHMEAVKVYIHRDLKSRNV 110
DB	208	LTFELPRGSLYRLH--RPNQLDEKRMALDVAKGMVYLIHSHPT--VHRDLKSPNLL 265
QY	111	IADGVILKICDGRASRPHNT--THMSLVGTFFWMAPEYIQSLPVSETCDTVSYGVLM 168
DB	266	VDRKVVAVKCDVFLSRMKHTYLSKSTAGTPEWMAPEVLRNPNAREKDDVYSFGVLM 325
QY	169	MLTRVPEFKLEGLQVAMLVENKERTLTPSSCPBSFAELHQQWADAKKRPSSFOIIS 228
DB	326	LATSRLPVMGLNMQVGVAGFONRLRLEIPDDIDLTLVAQITRECWTEPHLRSPFTLMO 385
QY	229	ILESMSNTSLPDKNS 245
DB	386	SLKRLQG-LNISRANT 401

##### RESULT 2

JU0229

mixed-lineage protein kinase 1 - human

C/Species: Homo sapiens (man)

C/Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #ext\_change 17-Mar-1999

C/Accession: S12467; JU0229

R/Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.

Eur. J. Biochem 213, 701-710, 1993

A/Title: Identification of a new family of human epithelial protein kinases containing:

A:Reference number: S32467; MUID:93238756; PMID:8477742  
A:Accession: S32467  
A:Molecule type: mRNA  
A:Residues: 1-394 <DO>  
C:Genetics:  
A:Gene: GDB:MLK1  
A:Cross-references: GDB:141921; OMIM:600136  
A:Map position: 14q24.3-14q31  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein K  
F:1-269/Domain: catalytic <CAT>  
F:1-268/Domain: catalytic <CAT>  
F:9-17/Region: protein kinase ATP-binding motif  
F:289-310/Region: leucine zipper motif  
F:324-345/Region: leucine zipper motif  
F:354-368/Region: basic

Query Match 39.3%; Score 510.5; DB 2; Length 394;  
Best Local Similarity 42.7%; Pred. No. 2.1e-23;  
Matches 108; Conservative 44; Mismatches 76; Indels 25; Gaps 4;

1 YRAKMSODKEVAVK-----KILKIEKEAELISVLSHNIIOFYGVILLEPPNY 48  
18 YRAFMIGD-EVAVKAAHDEDEDISOTIENVROGAKLFAVLKHPNIIALRGVCLKEPNTL 75  
49 GIVTEYASLSLYDINSNRSEMDHMTWADVAKGHIYLMKAPVYKTHRDLSKR 108  
76 CLVMEFARGGLINRVLISKR--IPPDILVWVAQIAGMNYLHDEALVPIIHDLSKN 132  
109 VVI-----AAGVLKICDFGASRFPHNTTMSLVGTFPMMAPEVIOQLPVSECTDTY 160  
133 ILILQKVENGLSNKILKITDFGLAREHRTTKSAAGTYAMMAPEVIRASMSKGSQVW 192  
161 SYGVLMELTREVPEKGLGLOVAMLYVEKNERLTIPSSCFPSAEILHOCWEADAKR 220  
193 SYGVLMELTREVPEKGLGLOVAMLYVEKNERLTIPSTCEPAPKIMEDCWNDDPSHR 252  
221 PSFKQIISLSESM 233  
253 PSFTNLDQLTTI 265

RESULT 3  
116747  
Hypochemical protein R13F6.7 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 17-Mar-2000  
C:Accession: T16747  
A:Reference number: T16747  
A:Map position: 3  
A:Cross-references: GDB:141921; OMIM:600136  
A:Map position: 14q24.3-14q31  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein K  
F:1-269/Domain: catalytic <CAT>  
F:1-268/Domain: catalytic <CAT>  
F:9-17/Region: protein kinase ATP-binding motif  
F:289-310/Region: leucine zipper motif  
F:324-345/Region: leucine zipper motif  
F:354-368/Region: basic

Query Match 38.1%; Score 495.5; DB 2; Length 328;  
Best Local Similarity 41.0%; Pred. No. 1.3e-22;  
Matches 102; Conservative 48; Mismatches 86; Indels 13; Gaps 5;

1 YRAKMSOD---KEVAVKILKIEKEAELISVLSHNIIOFYGVILLEPPNY-VGIVTEYAS 56  
67 FSNWMTLPDSSQRTIAKQVLEKEAELISVLSHNIIOFYGVILLEPPNY-VGIVTEYAS 126  
57 LGSLLDYINSNRSEEM-----DMHMTWADVAKGHIYLMKAPVYKTHRDLSKR 109

127 KSLVDFHSESSQGFASGNSFDVYVKKASQIASGIQVLAHYADVTIIHRDLKSKNV 186  
110 VIAADGVLIKICDFGASRFPHNTTMSLVGTFPMMAPE-VIOQLPVSECTDTYSGVILM 167  
187 VLDKNIIVCKICDFGASRFPHNTTMSLVGTFPMMAPE-VIOQLPVSECTDTYSGVILM 246  
168 EMTREVPEKGLGLOVAMLYVEKNERLTIPSSCFPSAEILHOCWEADAKRPSFKQII 227  
247 EILSKVPEKGLGLOVAMLYVEKNERLTIPSSCFPSAEILHOCWEADAKRPSFKQII 306  
228 SILEMSND 236  
307 GELNRLAEN 315

RESULT 4  
568178  
mixed-lineage protein kinase 2 (EC 2.7.1.1) - human  
C:Species: Homo sapiens (man)  
C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 11-Jun-1999  
C:Accession: S68178; I38044; S32468  
R:Biochem. 234, 492-500, 1995  
Eur. J. Biochem. 234, 492-500, 1995  
A:Title: Complete nucleotide sequence, expression, and chromosomal localization of human  
A:Reference number: S68178; MUID:96128179; PMID:8536694  
A:Accession: S68178  
A:Molecule type: mRNA  
A:Residues: 1-954 <DO>  
A:Cross-references: EMBL:X90846; NID:971419; PIDN:CAA62351.1; PID:9971420  
R:Kato, M.; Hixal, M.; Sugimura, T.; Terada, M.  
Oncogene 10, 1447-1451, 1995  
A:Title: Cloning and characterization of MST, a novel (putative) serine/threonine kinase.  
A:Reference number: I38044; MUID:95249256; PMID:7731697  
A:Accession: I38044  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-461, 'A', 'V', '465-470', 'S', '472-806', 'R', '808-817', 'A', '819-954' <RES>  
A:Cross-references: EMBL:Z48615; NID:9758592; PIDN:CAA8531.1; PID:9758593  
R:Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.  
Eur. J. Biochem. 213, 701-710, 1993  
A:Title: Identification of a new family of human epithelial protein kinases containing  
A:Reference number: S32467; MUID:93238756; PMID:8477742  
A:Accession: S32467  
A:Molecule type: mRNA  
A:Residues: 244-464, 'AQAAGRQPHALML' <DO>  
C:Genetics:  
A:Gene: GDB:MLK2; GDB:MST  
A:Cross-references: GDB:362654; GDB:624810; OMIM:600137  
A:Map position: 19q13.1-19q13.2  
C:Superfamily: mixed-lineage protein kinase 2; protein kinase homolog; SH3 homology  
C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein K  
F:23-76/Domain: SH3 homology <SH3>  
F:96-364/Domain: protein kinase ATP-binding motif  
F:104-112/Region: protein kinase ATP-binding motif  
F:384-405/Region: leucine zipper motif  
F:419-440/Region: leucine zipper motif  
F:449-463/Region: basic  
F:125,145,222,224/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 37.7%; Score 490.5; DB 1; Length 954;  
Best Local Similarity 42.3%; Pred. No. 7.2e-22;  
Matches 113; Conservative 35; Mismatches 92; Indels 27; Gaps 6;

1 YRAKMSODKEVAVKILKIEK-----EAILSVLSHNIIOFYGVILLEPPNY 48  
113 YRALM--RGEVAVKARLDEPKDPAYTAEGVCEALFGALQHPNIIALRGACINPPL 170  
49 GIVTEYASLSLYDINSNRSEMDHMTWADVAKGHIYLMKAPVYKTHRDLSKR 107  
171 CLVMEYARGALSRVLAGRVP-----DHVILVWVAQIAGMNYLHDEALVPIIHDLSKSI 226  
108 NVVI-----AAGVLKICDFGASRFPHNTTMSLVGTFPMMAPEVIOQLPVSECTDT 159



Db 294 LKSTKSPGATVAAWMAPEVIRNEPESEKVDIWSFGVIMELLTGRIPIKDVSSAIW 353  
 QY 187 LVAKENELTIPSSCPSPFAELHQCWEADAKRPPSKOIIISLESNSNT-SLPDK 242  
 Db 354 GVGNSLHLPPVSSCPDGFKILLRCQWNSKRNRPSRQIHLHDISADVLSTPGE 410

## RESULT 8

A53800  
 mixed-lineage protein kinase (BC 2.7.1.-) 3 - human  
 N/Alternate names: protein kinase PTK; protein kinase SPRK  
 C/Species: Homo sapiens (man)  
 C/Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 10-Sep-1999  
 C/Accession: A53800; I58395  
 R/Gallo, K.A.; Mark, M.R.; Scadden, D.T.; Wang, Z.; Gu, Q.; Godowski, P.J.  
 J. Biol. Chem. 269, 15092-15100, 1994  
 A/Title: Identification and characterization of SPRK, a novel src-homology 3 domain-cont  
 A/Reference number: A53800; MUID:94253068; PMID:8155146  
 A/Accession: A53800  
 A/Status: preliminary  
 A/Molecule type: mRNA

A/Residues: 1-847 <GAL>  
 A/Cross-references: GB:U07747; NID:9464027; PIDN:AAA19647.1; PID:9464028  
 R/ing, Y.L.; Leung, I.W.; Heng, H.H.; Tsui, L.C.; Lasseam, N.J.  
 Oncogene 9, 1745-1750, 1994  
 A/Title: MLK-3: identification of a widely-expressed protein kinase bearing an SH3 domai  
 A/Reference number: I58395; MUID:94239754; PMID:8183572  
 A/Accession: I58395  
 A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA  
 A/Residues: 1-647 <RBS>  
 A/Cross-references: GB:U22976; NID:9488295; PIDN:AAA59859.1; PID:9488296  
 C/Genetics:  
 A/Gene: GDB:MLK3; PTK1; SPRK  
 A/Cross-references: GDB:134755; OMIM:600050  
 A/Map position: 11q13.1-11q13.3

C/Superfamily: mixed-lineage protein kinase 3; protein kinase homology; SH3 homology  
 C/Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein  
 F/48-100/Domain: SH3 homology <SH32>  
 F/115-383/Domain: protein kinase homology <KIN>  
 F/123-131/Region: protein kinase ATP-binding motif  
 F/403-424/Region: leucine zipper motif  
 F/438-453/Region: leucine zipper motif  
 F/468-482/Region: basic

Query Match 37.1%; Score 482.5; DB 1; Length 847;  
 Best Local Similarity 41.1%; Pred. No. 1.9e-21;  
 Matches 104; Conservative 44; Mismatches 80; Indels 25; Gaps 6;

QY 1 YRAKWTISQ-----DKAVAVKKLKIKEAEILSVLSHRNIIQFYGVILEPPNYG 49  
 Db 132 YRSMGELVAVKARODPEDISV-TAHSVQEARLFMLAHPIITALKAVCLEEPNIC 190  
 QY 50 IYTVYASLSGLYDINSNSEEMDMHIM-TMATDVAKGMHYLMEAPVKYIHRDLKSN 108  
 Db 191 LVMEIYAGGLSALAGRRP---PHVLVMAVOIARGHYIHGALVPIVHRDLKSN 246  
 QY 109 VV-----IADGV-----LKICDPSASSPHNHTTMSLVGTFPMAPVIOGLPYSEICDTY 160  
 Db 247 ILLQPIESDMEHKTITDFGLARBWHKTQMSAAGTYAWMAPIVAKSTFSKSDSW 306  
 QY 161 SYGVATLMTREVPFGLEGLOVAMLVYKNERLTISSCPSPFAELHQCWEADAKR 220  
 Db 307 SFGVTLMEILLTGVPKRGIDCLVAAGVAVNKLTLPISTCEPFAQLMADCAQDPHR 366  
 QY 221 PSFKOIIISLESN 233  
 Db 367 PDFASTIQGLEAL 379

RESULT 9  
 T48544

MAPK delta-1 protein kinase - Arabidopsis thaliana  
 N/Alternate names: protein kinase 20  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 20-Apr-2000 #sequence revision 20-Apr-2000 #text change 20-Apr-2000

C/Accession: T48544  
 R/Bevan, M.; Hilbert, H.; Braum, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft,  
 submitted to the Protein Sequence Database, April 2000  
 A/Reference number: 224490  
 A/Accession: T48544  
 A/Status: preliminary

A/Molecule type: DNA  
 A/Residues: 1-886 <BEV>  
 A/Cross-references: EMBL:AL163812  
 A/Experimental source: cultivar Columbia; BAC clone F14F18  
 C/Genetics:  
 A/Map position: 5  
 A/Intons: 148/3; 180/3; 327/3; 362/2; 568/3; 658/1; 671/3; 694/3; 728/2; 754/3; 810/3;  
 A/Note: F14F18.20

Query Match 36.9%; Score 480; DB 2; Length 886;  
 Best Local Similarity 42.9%; Pred. No. 2.8e-21;  
 Matches 97; Conservative 41; Mismatches 72; Indels 16; Gaps 5;

QY 1 YRAKWTISQDEKVAVKLL-----KIEKAEILSVLSHRNIIQFYGVILEPPNYG 50  
 Db 664 YRAEW--NGTEVAVKKFLDODFGDALTFQKSEIEMLRHPNVVLEFGAVTRPPNFSI 721  
 QY 51 VTVYASLSGLYDINSNSEEMDMHIM-TMATDVAKGMHYLMEAPVKYIHRDLKSNV 110  
 Db 722 LTFELPRGSLYRLH--RPNQIDKERRMALDVAKGMHYLHSHNT-VVHDLKSNLL 779  
 QY 111 IADGVKLICDPSASSPHNHT--THMSLVGTFPMAPVIOGLPYSEICDTYSYGVIME 168  
 Db 780 VDKMNVKVCDFGLSRMKHTYLSKSTAGTPEWMAPEVIRNEPESEKVDIWSFGVIME 839  
 QY 169 MLTRVDFPKLEGLQVAMLVYKNERLTISSCPSPFAELHQCWE 214  
 Db 840 LATSRVPMKGLMPQVAVGAVGFQNRLEIPDDIDTLVAQITRECWQ 885

## RESULT 10

F96763  
 hypothetical protein F25P22.8 (imported) - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 02-Mar-2001 #sequence revision 02-Mar-2001 #text change 31-Mar-2001  
 C/Accession: F96763  
 R/Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Huizart, L.  
 Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A86141; MUID:21016719; PMID:11130712  
 A/Accession: F96763  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-1030 <STO>  
 A/Cross-references: GB:AE005173; NID:96692730; PIDN:AAF24836.1; GSPDB:GN00141  
 C/Genetics:  
 A/Gene: F25P22.8  
 A/Map position: 1

Query Match 36.6%; Score 475.5; DB 2; Length 1030;  
 Best Local Similarity 38.3%; Pred. No. 5.9e-21;  
 Matches 102; Conservative 48; Mismatches 85; Indels 31; Gaps 7;

QY 1 YRAKWTISQDEKVAVKLL-----KIEKAEILSVLSHRNIIQFYGVILEPPNYG 50  
 Db 763 YRGDM--HGTVAVAKKFLDODLTGALAEFRSEVRIMKRLHPNIVLFMGAVTRPPNLSI 820

[illegible]

## RESULT 11

protein kinase 6 (EC 2.7.1.1) - soybean  
 C/Species: Glycine max (soybean)  
 C/Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 24-Sep-1999  
 C/Accession: S29851, S27760  
 R/Peng, X.H.; Zhao, Y.; Bottino, P.J.; Kung, S.  
 Blochim. Biophys. Acta 1172, 200-204, 1993  
 A/Title: Cloning and characterization of a novel member of protein kinase family from *Sesquipedalia*  
 A/Accession number: S29851, MUID:93176812; PMID:8439562  
 A/Accession: S29851  
 A/Molecule type: mRNA  
 A/Residues: 1-462 <P>  
 A/Cross-references: EMBL:ME7449; NID:G170046; PIDD:AA34002.1; PID:G170047  
 C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homoid  
 C/Keywords: ATP; phosphotransferase  
 F/154-419/Domain: protein kinase homology <KIN>

Query Match	35.4%	Score 460.5;	DB 2;	Length 462;
Best Local Similarity	37.0%	Pred. No. 2.1e-20;		
Matches	98;	Conservative	53;	Mismatches 81;
				Indels 33;
				Cross 7;

```

0Y 6 ISODEVAVKLL-----KIEK---EAMISVTSXHNITIOFYVILPEPNYG 49
Db 174 YKDEAVAVKIIWVEDDONGALASRLKOPFIREVTLTISRHHQWIVFSAACKRPYIC 233
0Y 50 IYTBYSIGSLVDYINSRSEBMDHIMTATDVAKGMHYLHMEAPKVBHRLDKSRNV 109
Db 234 IITBYLAEIGSRAYLHKLEHOTISLOKLIARFALDIAREMEYTHSQG---YIHRULKENI 299
0Y 110 VIAAGVLIKIDPFGASRFHNHTTMSLV---GTFPMWAPREVISOFLPVSETCDTYSYGV 165
Db 291 LINEDNHKIIDPGLA---CEBASCDLLADPGYIRWMAPEMIRKRSYGKKVDYVSFGLI 347
0Y 166 LWMELTREVPKDEGLQVAVLVEKNERLTI PSSCPSPFELLHQCEADAKKRPBFQ 225
Db 348 LWMELTGFIPEYDNMPIOAPAVVNVKNSRPIIPBNCPPARBALIEQCMSLOPDKRBEFW 407
0Y 226 IISILE-----SMSNDTSL--PDKC 243
Db 408 VKLILQEPSSLASDGTSLSLVNPFC 432

```

## RESULT 12

Protein kinase homolog F2401.13 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 23-Mar-2001  
C/Accession: T01451  
R/Shim, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Cor-  
eoligis, A.; Ecker, J.R.  
submitted to the EMBL Data Library, January 1998  
A/Description: Genomic sequence for Arabidopsis thaliana BAC F2401.  
A/Reference number: Z14211  
A/Accession: T01451

A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-390 <SH1>  
A:Cross-references: EMBL:AC003113; NID:g2e69438; PID:g2781357; GSPDB:GN00055; ATSP:F240  
C:Genetics:  
A:Gene: ATSP:F2401.13  
A:Map position: 1  
A:Introns: 149/3; 301/3  
C:Superfamily: kinase-related transforming protein; protein kinase homology

Query Match	35.2%	Score 457.5;	DB 2;	Length 390;
Best Local Similarity	37.2%;	Pred. No. 2.7e-20;		
Matches 96;	Conservative 49;	Mismatches 86;	Indels 27;	Gaps 6;

Query Match	35.2%	Score 457.5;	DB 2;	Length 390;
Best Local Similarity	37.2%	Pred NC 3	7a-20.	

Qy	1	YRAKMTSODKEVAAYKTL-----LKIER----	EAEILSVLSHRNIIQOFVILLEPP	46
Db	101	YRG--YKQRAAAYKVKVR1PTHKETRAKLEQCKESEVALLSRLEPHNIQVIAACKRP	158	
Qy	47	NYGIVTEVASLGSIVDYIYNSNRSEEMDDHMTMATVAQGMHYLHNEAPVKYTHDDLKS	106	
Db	159	VYCIIEYMSQGNLMTYLNKKEPEYSLSLEYVLRALDLSRGMELYSQG---VYHRDLKS	215	
Qy	107	RNVVIAADGVLCIKCPFGASRFPHNHTH--MSLVGTFPMAABEVIQSLPVSETCDYYSYGV	165	
Db	216	NNLLNLDENRVRKADPGTSCLETCREAKGMNGYTRMNAEMIKERPYRKVDVYSFGIV	275	
Qy	166	LWMLTREVPEFGLEGLVVAWLVEYKNEBRLTIPSSCRSPAEILHHCWEADAKRPSFQ	225	
Db	276	LWELTALLPFGGMPTVQAFAVAEKNERPPLPASCCOPALAHILIKRCWSENSKQPDFSN	335	
Qy	226	ISILIESMNDTSLPDKC 243		
Db	336	IYAAVLEKY-----DEC 346		

### RESULT 13

probable protein kinase TCRR2 - tomato  
C:Species: Lycopersicon esculentum (tomato)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 08-Oct-1999  
C:Accession: T06576  
R:Hackett, R.M.  
submitted to the EMBL Data Library, March 1998  
A:Reference number: Z15770  
A:Accession: T06576  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-982 <HAC>  
A:Cross-references: EMBL:AJ005077; NID:e1296722; PIDN:CA06334.1; PID:e12967233  
A:Experimental source: cultivar Alisa Craig  
C:Genetics:  
A:Gene: TCRR2

Query Match	34.8%	Score 453;	DB 2;	Length 982;
Best Local Similarity	38.5%	Pred. No. 1.2e-19;		
Matches 101; Conservative	45.1%	Mismatch 10		

```

27 1 KQNMSSVQAIAVAKLL-----KIEKEKELISVSHNIIQFGLLEPPNGI 50
Dd 716 YHADM--NGTEVAVAKKFLDQDFSGALAEFRERIRLRHHNVVRFGALTTRPHHSI 77
Qy 51 VHEVYALSLGYIYINSNSSEMDMDHMTATVAGMHYLHNEAEVKYIHRDLKSRNV 11
Dd 774 IIEFLPRGSLYRIIHRPHF-QIDEROKIKKALVAVAGMCIHTSNPT-IVHRDLKSPLL 83
Qy 111 IADGVLTICDFGASRPFNNHT--THMSLVGTFFPMAPEVIOSLPSESTDYISYGVILME 16
Dd 832 VPTDMNVKICDFGSLRKANTFTLSGKSTACTPEPMAPEVLRNEPSNEKDIYSFVILME 89
Qy 169 MLTRVEPKGLEGHQVAMLVENKNERLTTPSCPSRFAELHQCWEADAKKRPKQIIS 22
Dd 892 LATLRLPMSGNNPQVGVAGFQNKRLIEPKELDPVIAIIMECQDTPNLRPSAQLTV 95

```

QY 229 ILISMSNDTSLP---DKCSFL 247  
 Db 952 ALFPLQR-LVIFPAVDQNSRL 972

## RESULT 14

T10671  
 protein kinase homolog F6E21.90 - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cross)  
 C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #ext\_change 19-May-2000  
 C/Accession: T10671  
 R/Bevan, M.; Lemard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro  
 submitted to the Protein Sequence Database, June 1999  
 A/Accession: T10671  
 A/Molecule type: DNA  
 A/Residues: 1-412 <BEV>  
 A/Cross-references: EMBL:AL049914; GSPDB:GN00062; ATSP:F6E21.90  
 A/Experimental source: cultivar Columbia; BAC clone F6E21  
 Genetics:  
 Gene: ATSP:F6E21.90  
 Map position: 4  
 A/Intons: 300/2  
 C/Superfamily: kinase-related transforming protein; protein kinase homology  
 F:135-392/Domain: protein kinase homology <KIN>

Query Match 34.5%; Score 448; DB 2; Length 412;  
 Best local similarity 36.3%; Pred. No. 1e-19;  
 Matches 91; Conservative 56; Mismatches 84; Indels 20; Gaps 4;

QY 1 YRAKMIQDPKEVAVKLT-----LKIKKAEIISVLSHRNIQFYGVILEPP 46  
 Db 146 YRGTYNGED--VAIKLERSDSNPKEAQALFQOEVSMLAFLEHGPVIVFVIGAKLPEM 203  
 QY 47 NGYITVEVASLSGLDYIN-SNRSEMDMDHIMTVAITDAKGMHYLHMEAPVVIHRDLKS 106  
 Db 204 VMCIVTEVAKGSGVRFILTRKQRAVPLKLAVMQALDVAKGMAYHER---NFIHRDLKS 260  
 QY 107 RNVVIAADGVLIKICPGASRFNHTHMS-LVGTPEPMAPEVIGSLPVSECTDYISYGV 165  
 Db 261 DNLISADSKSIKADFGVARIEVQTEGMPETGYRWMAPEMIOHRPYQKDVVSFGIV 320  
 QY 166 LWMELTREVPEFKLEGLOVAVLVEKNERLITPSSCPSPFALLHOCWEADAKKRPSPFO 225  
 Db 321 LWEILITGLLPFQMTAVOAFVAVNRGVPTVPADCLPVLGEIMTRCWDADPEVAPCPAE 380  
 QY 226 ILISMSND 236  
 381 LVNLEAAETG 391

## RESULT 15

T48400  
 serine/threonine-protein kinase ctrl - Arabidopsis thaliana  
 N/Alternate names: protein F17C15.150  
 C/Species: Arabidopsis thaliana (mouse-ear cross)  
 C/Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #ext\_change 17-Nov-2000  
 C/Accession: T48400; A45178  
 R/Bevan, M.; Pohl, T.; Weizengger, T.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.;  
 submitted to the Protein Sequence Database, March 2000  
 A/Reference number: Z24492  
 A/Accession: T48400  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-821 <BEV>  
 A/Cross-references: EMBL:AL162506  
 A/Experimental source: cultivar Columbia; BAC clone F17C15  
 R/Kieber, J.U.; Rothenberg, M.; Roman, G.; Feldmann, K.A.; Ecker, J.R.  
 Cell 72, 427-441, 1993  
 A/Title: CK1, a negative regulator of the ethylene response pathway in Arabidopsis, enc  
 A/Reference number: A45178; MUID:93161417; PMID:8431946  
 A/Contents: Columbia  
 A/Accession: A45178

A/Status: preliminary; not compared with conceptual translation  
 A/Molecule type: DNA  
 A/Residues: 1-468,470-821 <KIB>  
 A/Note: sequence extracted from NCBI backbone (NCBIP:124878)  
 C/Genetics:  
 A/Map position: 5  
 A/Intons: 216/3; 317/2; 364/3; 399/2; 468/1; 560/1; 573/3; 596/3; 630/2; 658/3; 691/3;  
 A/Note: F17C15.150  
 C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol  
 C/Keywords: AtP  
 F:549-812/Domain: protein kinase homology <KIN>  
 F:557-565/Region: protein kinase ATP-binding motif

Query Match 34.0%; Score 442; DB 2; Length 821;  
 Best local similarity 40.2%; Pred. No. 4.5e-19;  
 Matches 99; Conservative 41; Mismatches 90; Indels 16; Gaps 5;

QY 1 YRAKMIQDPKEVAVKLT-----KAEIISVLSHRNIQFYGVILEPPNGI 50  
 Db 566 HRAEMHGSQ--VAVKIMEQDPHAEVNEFLREVAIMKRLNHPNIVLFMGAVTQPPNLSI 623  
 QY 51 VTEYASLSGLDYIN-SNRSEMDMDHIMTVAITDAKGMHYLHMEAPVVIHRDLKSRNV 109  
 Db 624 VTEYLSRSGSLYRLHLKSGAREQLDERRRLSMAYDVAKGMHYLHMRNP-PIYHRDLKSPNL 682  
 QY 110 VIAADGVLIKICPGASRFNHT--THMSLVGTPEPMAPEVIGSLPVSECTDYISYGVILM 167  
 Db 683 LVNKKITVACDFGLSRKLAFTLSKSAAGTPENWAPAEVLRDEPSNKSQDVYSFGVILM 742  
 QY 168 EMLTREVPEFKLEGLOVAVLVEKNERLITPSSCPSPFALLHOCWEADAKKRPSPFO 227  
 Db 743 ELATLQOPWGNLNPAGVAVVAVGFKCKRLIIPRNLPQVAALIEGCTWEPKRPSPFALIM 802  
 QY 228 ILISMSND 233  
 Db 803 DLRLPL 808

Search completed: May 1, 2003, 20:38:39  
 Job time: 16.8131 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:13:36 ; Search time 12.4017 Seconds.

(without alignments)  
826,070 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_31\_277

Sequence: 1 YRAKMSIQDEKVAVKLLKI.....SLESMSNDTSLPDKNSFL 247

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5 ;

Searched: 112892 seqs, 41476328 residues

al number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwIsaProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	510.5	39.3	394	1 M3K9_HUMAN	P01012 homo sapien
2	490.5	37.7	954	1 M3KC_HUMAN	Q02779 homo sapien
3	490	37.7	859	1 M3KC_HUMAN	Q12852 homo sapien
4	490	37.7	888	1 M3KC_MOUSE	Q07000 mus musculu
5	488	37.5	888	1 M3KC_RAT	Q05609 rattus norv
6	442	34.0	821	1 CTR1_ARATH	Q62073 mus musculu
7	395	30.4	579	1 M3K7_MOUSE	Q43318 homo sapien
8	395	30.4	606	1 M3K7_HUMAN	P00519 homo sapien
9	369.5	28.4	1130	1 ABL1_HUMAN	P00521 abelson mur
10	368.5	28.3	746	1 ABL1_MLVB	P00520 mus musculu
11	368.5	28.3	1123	1 ABL1_HUMAN	P02684 homo sapien
12	365.5	28.1	1182	1 ABL1_MOUSE	P00522 drosophila
13	364.5	28.0	1520	1 ABL1_DROME	P10447 feline barc
14	359.5	27.3	1584	1 ABL1_MOUSE	P04477 feline barc
15	351.5	27.0	1439	1 ABL1_MOUSE	P04477 feline barc
16	349.5	26.9	393	1 M3K7_DROME	P03104 drosophila
17	345	26.5	536	1 FYN_XIPHE	P27446 xiphophorus
18	343	26.4	505	1 FYN_HUMAN	P26655 homo sapien
19	337.5	26.0	536	1 FYN_XENLA	P06241 xenopus lae
20	335.5	25.8	536	1 FYN_HUMAN	P21804 gallus galli
21	335	25.8	822	1 FGRI_CHICK	P11362 homo sapien
22	335	25.8	822	1 FGRI_HUMAN	P16092 mus musculu
23	335	25.8	822	1 FGRI_MOUSE	P39688 mus musculu
24	334.5	25.7	533	1 FYN_MOUSE	P15054 avian sarco
25	332.5	25.5	587	1 SRC_AVIS2	P14234 mus musculu
26	332	25.5	517	1 FGR_MOUSE	P18151 dicyostell
27	330	25.4	410	1 KYK2_DICDI	Q04589 rattus norv
28	330	25.4	822	1 FGRI_RAT	P03949 caenorhabd
29	329.5	25.3	1196	1 ABL1_CAEL	P18460 gallus galli
30	329	25.3	806	1 CEK2_CHICK	P11240 homo sapien
31	328.5	25.3	450	1 CSK_HUMAN	P00523 gallus galli
32	328.5	25.3	532	1 SRC_CHICK	P13115 xenopus lae
33	327.5	25.2	531	1 SRC1_XENLA	

34	327.5	25.2	531	1 SRC2_XENLA	P13116 xenopus lae
35	327	25.2	806	1 FGR_HUMAN	P22607 homo sapien
36	326.5	25.1	557	1 SRC_AVIS2	P14085 avian sarco
37	325.5	25.0	533	1 FYN_CHICK	Q05876 gallus galli
38	324.5	25.0	535	1 SRC_HUMAN	P12931 homo sapien
39	324.5	25.0	568	1 SRC_AVIS3	P14084 avian sarco
40	324.5	25.0	801	1 FGR3_MOUSE	Q61851 mus musculu
41	324	24.9	535	1 YRK_MOUSE	Q02977 gallus galli
42	323	24.8	812	1 FGRI_CHICK	P22182 xenopus lae
43	322.5	24.8	450	1 CSK_RAT	P25777 rattus norv
44	321	24.7	528	1 YES_AVISY	P00527 avian sarco
45	321	24.7	541	1 YES_CHICK	P09324 gallus galli

## ALIGNMENTS

RESULT 1  
M3K9\_HUMAN STANDARD; PRT; 394 AA.  
ID M3K9\_HUMAN  
AC P01012;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Mitogen-activated protein kinase kinase 9 (EC 2.7.1.-) (Mixed  
DE lineage kinase 1) (Fragment).  
GN MAP3K9 OR MLK1 OR PRK1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NBI; taxid:9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Colon epithelium;  
RX MEDLINE=93238756; Pubmed=8477742;  
RA Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;  
RT "Identification of a new family of human epithelial protein kinases  
RT containing two leucine/isoleucine-zipper domains.";  
RL Eur. J. Biochem. 213:701-710(1993).  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN EPITHELIAL TUMOR CELL LINES OF  
CC COLONIC, BREAST AND OESOPHAGEAL ORIGIN.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC MAP KINASE KINASE KINASE SUBFAMILY.  
CC PIR; S32467; S32467.  
DR PIR; J00229; J00229.  
DR HSSP; P12931; 1FMK.  
DR Genew; HGNC:6861; MAP3K9.  
DR MIM; 600136;  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR Prodom; PPO00001; Euk\_pkinase; 1.  
DR SMART; SM00219; Tyrc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PSS0011; PROTEIN\_KINASE\_DOM; 1.  
KM Transferrase; Serine/threonine-protein kinase; Tyrosine-protein kinase;  
KW ATP-binding.  
FT NON\_TER 1 1  
FT DOMAIN 3 271 PROTEIN KINASE.  
FT NE\_BIND 9 17 ATP (BY SIMILARITY).  
FT BINDING 30 30 ATP (BY SIMILARITY).  
FT ACT\_SITE 127 127 BY SIMILARITY.  
FT DOMAIN 289 310 LEUCINE-ZIPPER 1 (BY SIMILARITY).  
FT DOMAIN 324 345 LEUCINE-ZIPPER 2 (BY SIMILARITY).  
FT DOMAIN 354 368 ARG/LYS-RICH (BASIS).  
SQ SEQUENCE 394 AA; 44975 MW; DBE40B7D31047FD8 CRC64;  
Query Match 39.3%; Score 510.5; DB 1; Length 394;  
Best Local Similarity 42.7%; Pred. No. 56-34;  
Matches 108; Conservative 44; Mismatches 76; Indels 25; Gaps 4;

QY 1 YRAWMISODKEVAVK-----KLKIEKEAEILSVLSHRNIIQFYGVILEPPNY 48  
 DB 18 YRAWMIGD--EVAVKARHDEDEDSOTIENROEAKLPAULKPHNIIALRGVCLKEPNTL 75  
 QY 49 GIVTEYASISGLYDYNINSEEMDMHIMTADTVAKMHEVILHMEAPVKVITHRDLKSRN 108  
 DB 76 CLVMEFARGGLNVLGSKR---IPDLVAMVAQIARGMNYLDEALVPIIHRDLKSSN 132  
 QY 109 VVI-----ADGVLIKICDFGASRFNHTTMSLVGTFPMMAPEVIOQLPVSCTDPT 160  
 DB 133 ILILQKVENGLSNKILKITDFGLAREWHRTTKMSAGTYAMMAPEVIRASMFSGKSPVW 192  
 QY 161 SYGVVLMEMLTREVPFKELBGLQVAMLVVEKNERLTIPSSCPSPFAELLHQCWEADAKK 220  
 DB 193 SYGVLLMELLTGEVPRFGIDGLRVAVGVAMNKLALPISTCEPPAKIMEDCMNDPDRSR 252  
 QY 221 PSFKQIISILESM 233  
 DB 253 PSFTNILDQLTTI 265

-----  
 M3KA HUMAN STANDARD; PRT; 954 AA.  
 ID M3KA HUMAN STANDARD; PRT; 954 AA.  
 AC 002779; 012761; 014871;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Mitogen-activated protein kinase kinase kinase 10 (EC 2.7.1.37)  
 GN MAP3K10 OR MKK2 OR MST.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=96128179; PubMed=8536694;  
 RA Dorew D.S., Devereux L., Tu G.F., Price G., Nicholl J.K.,  
 RA Sutherland G.R., Simpson R.J.;  
 RT "Complete nucleotide sequence, expression, and chromosomal  
 RT localisation of human mixed-lineage kinase 2.",  
 RT Eur. J. Biochem. 234:492-500(1995).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Brain;  
 RC TISSUE=Brain;  
 RX MEDLINE=95249256; PubMed=7731697;  
 RA Katoh M., Hirai M., Sugimura T., Terada M.;  
 RT "Cloning and characterization of MST, a novel (putative)  
 RT serine/threonine kinase with SH3 domain.",  
 RT Oncogene 10:1447-1451(1995).  
 RL [3]  
 RN SEQUENCE OF 244-480 FROM N.A.  
 RP TISSUE=Colon epithelium;  
 RC MEDLINE=93238756; PubMed=8477742;  
 RA Dorew D.S., Devereux L., Dietzsch E., de Kretser T.,  
 RT "Identification of a new family of human epithelial protein kinases  
 RT containing two leucine/isoleucine-zipper domains.",  
 RT Eur. J. Biochem. 213:701-710(1993).  
 RL [4]  
 RN CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND SKELETAL MUSCLE.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MAP KINASE KINASE KINASE SUBFAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; X90846; CAAG2351.1; --  
 DR EMBL; Z40615; CAAG8531.1; --  
 DR PIR; S32468; S32468.  
 DR HSSP; P11362; 1FGK.  
 DR Genew; HGNC:6849; MAP3K10.  
 DR MIM; 600137; --  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00018; SH3; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SM00221; STYKc; 1.  
 DR SMART; SM00221; STYKc; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE\_ST; 1.  
 DR PROSITE; PS50011; PROTEIN KINASE\_DOM; 1.  
 DR PROSITE; PS50002; SH3; 1.  
 DR Transferrase; Serine/threonine-protein kinase; Tyrosine-protein kinase;  
 KW ATP-binding; SH3 domain.  
 FT DOMAIN 2  
 FT DOMAIN 16  
 FT DOMAIN 81  
 FT NP\_BIND 98 360  
 FT NP\_BIND 104 112  
 FT BINDING 125 125  
 FT ACT\_SITE 222 222  
 FT DOMAIN 384 405  
 FT DOMAIN 419 440  
 FT DOMAIN 449 463  
 FT DOMAIN 462 464  
 FT CONFLICT 465 480  
 FT CONFLICT 471 471  
 FT CONFLICT 807 807  
 FT CONFLICT 818 818  
 SQ SEQUENCE 954 AA; 103623 MW; 538F4AA559B0ABA CRC64;

Query Match 37.7%; Score 490.5; DB 1; Length 954;  
 Best Local Similarity 42.3%; Pred. No. 5.7e-32; Indels 27; Gaps 6;  
 Matches 113; Conservative 35; Mismatches 92;  
 QY 1 YRAWMISODKEVAVKKL-LKIEK-----EAEILSVLSHRNIIQFYGVILEPPNY 48  
 DB 113 YRAWM--RGEVAVKARLPDENOPAVTARGVCEARLFGALQHPNIIALRGACTLNPPL 170  
 QY 49 GIVTEYASISGLYDYNINSEEMDMHIM-TAATDVAKMHEVILHMEAPVKVITHRDLKSR 107  
 DB 171 CLVMEFARGGLNVLGSKR---PHVLVAMVAQIARGMNYLDEALVPIIHRDLKSI 226  
 QY 108 NVVI-----ADGVLIKICDFGASRFNHTTMSLVGTFPMMAPEVIOQLPVSCTDPT 159  
 DB 227 NILLLEAIENNLNDVTLKTDPLGLAREWHRTTKMSAGTYAMMAPEVIRASMFSGKSDV 286  
 QY 160 YSYGVVLMEMLTREVPFKELBGLQVAMLVVEKNERLTIPSSCPSPFAELLHQCWEADAKK 219  
 DB 287 MSFEVLLMELLTGEVPRFGIDGLRVAVGVAMNKLALPISTCEPPAKIMEDCMNDPDRPHG 346  
 QY 220 PSFKQIISILESMNDTSLPDKCNF 246  
 DB 347 RPDGSIILKRLVIEOGALFQMPLESF 373

-----  
 RESULT 3  
 M3KA HUMAN STANDARD; PRT; 859 AA.  
 ID M3KA HUMAN STANDARD; PRT; 859 AA.  
 AC 012852;  
 DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Mitogen-activated protein kinase kinase 12 (EC 2.7.1.37)  
 GN MAP3K12 OR ZPK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Teratocarcinoma;  
 RX MEDLINE=94311945; PubMed=8037767;  
 RA Reddy U.R., Pleasure D.;  
 RT "Cloning of a novel putative protein kinase having a leucine zipper  
 domain from human brain."  
 RL Biochem. Biophys. Res. Commun. 202:613-620(1994).  
 CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.  
 CC Phosphorylates beta-casein, histone 1 and myelin basic protein in  
 CC vitro.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -1- COFACTOR: Magnesium.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By  
 CC similarity).  
 CC -1- TISSUE SPECIFICITY: Highly expressed in brain and kidney.  
 CC -1- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol  
 CC under basal conditions and dephosphorylated when membrane-  
 CC associated (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MAP KINASE KINASE KINASE SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR HSBP; U07358; AAA67343.1; -  
 DR HSBP; P12931; IFMK.  
 DR Genew; HGNC:6851; MAP3K12.  
 DR MIM; 600447; -  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR Prodom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00221; STYKC.1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR Transelase; Serine/threonine-protein kinase; ATP-binding;  
 KM Phosphorylation; Magnesium; Membrane.  
 FT DOMAIN 125 366 PROTEIN\_KINASE.  
 FT NP\_BIND 131 139 ATP (By SIMILARITY).  
 FT BINDING 152 152 ATP (By SIMILARITY).  
 FT ACT\_SITE 236 236 BY SIMILARITY.  
 FT DOMAIN 665 668 POLY-PRO.  
 FT DOMAIN 720 725 POLY-GLU.  
 SO SEQUENCE 859 AA; 93188 MW; 0E5209792C5C6F05 CRC64;  
 Query Match 37.7%; Score 490; DB 1; Length 859;  
 Best Local Similarity 44.3%; Pred. No. 5.6e-32;  
 Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;  
 QY 10 KEVAVKLLKIEKAEI--LSVSHRNIOHYVILRPNGIYTEVASIGSLYDIYNSN 67  
 DB 147 EBAVAVKVRDI-KETDKLHRLKLGHPNITIFKGVCTAPCTCIMECAQSKQLEVARAG 205  
 QY 68 RSEEMDMHITVAATDVAKGNGVYLHMEAPVYVIRHDLKSRNVVIAAGVGLICDFGSR- 126  
 DB 206 RPTVPSL-LVDNEMGIAGGNVYLH---KIHRDLKSPMLITVDVYKISDFGSK 260

QY 127 FHNHTHSLVGFPPMAPEVIOSLPVSEICDTYSVGVLMMETREVPKGLGLOVAM 186  
 DB 261 LSDSTKMSFPGIVAMAPVETIRNEPVSEKVIDMSFVWMLTGBIPKDVDSAIW 320  
 QY 187 LVYKNERLTIPTSSCFPAELLHQWEADAKKRSFKQIISLESNSNT-SLPDK 242  
 DB 321 GVGNSLHLPVPSSCPDGFILRLQCNKSNRPNRFRQLHLHDIASADVLSPQE 377  
 RESULT 4  
 M3KC MOUSE STANDARD; PRT; 888 AA.  
 AC 060700; P70286;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Mitogen-activated protein kinase kinase 12 (EC 2.7.1.37)  
 DE (leucine-zipper protein kinase) (ZPK) (Dual leucine zipper bearing  
 DE kinase) (DLK).  
 GN MAP3K12 OR ZPK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CD-1; TISSUE=Brain;  
 RX MEDLINE=95074107; PubMed=7983011;  
 RA Holzman L.B., Merritt S.E., Fan G.;  
 RT "Identification, molecular cloning, and characterization of dual  
 RT leucine zipper bearing kinase. A novel serine/threonine protein kinase  
 RT that defines a second subfamily of mixed lineage kinases."  
 RT J. Biol. Chem. 269:30808-30817(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1CR X Swiss Webster; TISSUE=Brain;  
 RX MEDLINE=96365388; PubMed=8769565;  
 RA Blouin R., Beaudoin J., Bergeron P., Nadeau A., Grondin G.;  
 RT "Cell-specific expression of the ZPK gene in adult mouse tissues."  
 RT DNA Cell Biol. 15:631-642(1996).  
 RN [3]  
 RP PHOSPHORYLATION, AND MUTAGENESIS OF LYS-185 AND GLU-192.  
 RX MEDLINE=96279269; PubMed=8663324;  
 RA Mata M., Merritt S.E., Fan G., Yu G.G., Holzman L.B.;  
 RT "Characterization of dual leucine zipper-bearing kinase, a mixed  
 RT lineage kinase present in synaptic terminals whose phosphorylation  
 RT state is regulated by membrane depolarization via calcineurin."  
 RT J. Biol. Chem. 271:16888-16896(1996).  
 CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.  
 CC Phosphorylates beta-casein, histone 1 and myelin basic protein in  
 CC vitro.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -1- COFACTOR: Magnesium.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.  
 CC -1- TISSUE SPECIFICITY: Expressed in brain, kidney, lung, heart,  
 CC testis, gastrointestinal tract, stomach, liver and pancreas.  
 CC Within the nervous system, predominantly expressed in neurons and  
 CC enriched in synaptic terminals.  
 CC -1- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol  
 CC under basal conditions and dephosphorylated when membrane-  
 CC associated.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MAP KINASE KINASE KINASE SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: U14636; AAA57280.1; -  
DR EMBL: U23789; AAB17123.1; -  
DR HSSP: P12931; 1FMK.  
DR MGD; MG11346881; Mapk12.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR004040; STY\_pkinase.  
DR InterPro: IPR002290; Ser\_thr\_pkinase.  
DR Pfam: PF00069; pkinase; 1.  
DR Prodom: PD000001; Euk\_pkinase; 1.  
DR SMART: SM00221; STYKc; 1.  
DR PROSITE: PS00107; PROTEIN KINASE ATP; FALSE\_NEG.  
DR PROSITE: PS00108; PROTEIN KINASE ST; 1.  
DR TRANSFERASE; Serine/threonine-protein kinase; ATP-binding;  
KW Phosphorylation; Magnesium; Membrane.  
FT DOMAIN 158 399  
FT NP\_BIND 164 172  
FT BINDING 185 185  
FT ACT\_SITE 269 269  
FT DOMAIN 56 62  
FT DOMAIN 668 701  
FT DOMAIN 753 758  
FT MOTIFAGEN 185 185  
FT MOTIFAGEN 192 192  
FT CONFLICT 18 18  
FT CONFLICT 28 29  
FT CONFLICT 382 382  
FT CONFLICT 494 495  
FT CONFLICT 517 517  
FT CONFLICT 794 794  
SQ SEQUENCE 888 AA; 96083 MW; CFECLD34F889AB8 CRC64;  
Query Match 37.7%; Score 490; DB 1; Length 888;  
Best local Similarity 44.3%; Pred. No. 5.8e-32;  
Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;  
QY 10 KEAVVKLLIKKEAEI--LSVLSHRNIIQFYGVILEPPNYGIVTEYASISLYDIYNSN 67  
DB 180 EEVAVKVRDI-KETDIKHLKRLKHPNIIIFKGVCTQAPCYCLIMEFCAGQGLYEVLAR 238  
QY 68 RSEEMDMHMTATADVAKGMYLMEAPVYVIRHDLKSRNVVIAADGVLIKDFGASR- 126  
DB 239 RPTVPSL--LVDMGMIAGKMYLHLH--KIIRHDLSPNMLITDYDVYKISDFGSK 293  
QY 127 FHNHTTMSLVGTFFPMAPEVYIOSLPVSECTDYTSYGVVLWEMLTREVPPKGLQVAM 186  
DB 294 LSDSKTMSFPGTVAMAPVIRNEPVEKVDIWSFGVLMELTGLIPIKDVDSALIW 353  
DB 187 LVKKNERLITPSSCPSPFAELHQCWEADAKKRPSEKQIISLESMSNDT-SLPDK 242  
DB 354 GVGNSNHLHPVSSCPDGFKILRLQCMNSKPNRPSRQILHLHDIASADVLSTPQE 410  
RESULT 5  
MK3C\_RAT STANDARD; PRT; 888 AA.  
AC Q63796;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DB Mitogen-activated protein kinase kinase 12 (EC 2.7.1.37)  
DB (MAPK-upstream kinase) (MOK).  
GN MAPK12 OR MOK.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9622609; PubMed=8637721;  
RA Hirai S., Izawa M., Osada S., Spyrou G., Ohno S.;  
RT "Activation of the JNK pathway by distantly related protein kinases,

RT MEKK and MUK";  
RL Oncogene 12:641-650(1996).  
CC -!- FUNCTION: May be an activator of the JNK/SAPK pathway.  
CC Phosphorylates beta-casein, histone 1 and myelin basic protein in  
CC vitro.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC -!- COFACTOR: Magnesium.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By  
CC similarity).  
CC -!- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol  
CC under basal conditions and dephosphorylated when membrane-  
CC associated (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC MAP KINASE KINASE SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: D49785; BAA08621.1; -  
DR HSSP: P12931; 1FMK.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR004040; STY\_pkinase.  
DR InterPro: IPR002290; Ser\_thr\_pkinase.  
DR Pfam: PF00069; pkinase; 1.  
DR Prodom: PD000001; Euk\_pkinase; 1.  
DR SMART: SM00221; STYKc; 1.  
DR PROSITE: PS00107; PROTEIN KINASE ATP; FALSE\_NEG.  
DR PROSITE: PS00108; PROTEIN KINASE ST; 1.  
DR TRANSFERASE; Serine/threonine-protein kinase; ATP-binding;  
KW Phosphorylation; Magnesium; Membrane.  
FT DOMAIN 158 399  
FT NP\_BIND 164 172  
FT BINDING 185 185  
FT ACT\_SITE 269 269  
FT DOMAIN 56 62  
FT DOMAIN 668 671  
FT DOMAIN 698 701  
FT DOMAIN 753 758  
SQ SEQUENCE 888 AA; 96307 MW; 52AD96406BAE149 CRC64;  
Query Match 37.5%; Score 488; DB 1; Length 888;  
Best local Similarity 44.3%; Pred. No. 8.4e-32;  
Matches 105; Conservative 42; Mismatches 80; Indels 10; Gaps 6;  
QY 10 KEAVVKLLIKKEAEI--LSVLSHRNIIQFYGVILEPPNYGIVTEYASISLYDIYNSN 67  
DB 180 EEVAVKVRDI-KETDIKHLKRLKHPNIIIFKGVCTQAPCYCLIMEFCAGQGLYEVLAR 238  
QY 68 RSEEMDMHMTATADVAKGMYLMEAPVYVIRHDLKSRNVVIAADGVLIKDFGASR- 126  
DB 239 RPTVPSL--LVDMGMIAGKMYLHLH--KIIRHDLSPNMLITDYDVYKISDFGSK 293  
QY 127 FHNHTTMSLVGTFFPMAPEVYIOSLPVSECTDYTSYGVVLWEMLTREVPPKGLQVAM 186  
DB 294 LSDSKTMSFPGTVAMAPVIRNEPVEKVDIWSFGVLMELTGLIPIKDVDSALIW 353  
DB 187 LVKKNERLITPSSCPSPFAELHQCWEADAKKRPSEKQIISLESMSNDT-SLPDK 242  
DB 354 GVGNSNHLHPVSSCPDGFKILRLQCMNSKPNRPSRQILHLHDIASADVLSTPQE 410  
RESULT 6  
CTRL\_ARATH STANDARD; PRT; 821 AA.  
AC Q05609;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)

```

DR SMART; SMO0221; SYKC; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 65 69 POLY-GLY.
FT DOMAIN 135 141 POLY-GLY.
FT DOMAIN 551 809 PROTEIN KINASE.
FT NP_BIND 557 565 ATP (BY SIMILARITY).
FT BINDING 578 578 ATP (BY SIMILARITY).
FT ACT_SITE 676 676 E-Y-K; IN CTRL-4; EXHIBITS ETHYLENE-
FT MUTAGEN 596 596 E-Y-K; IN CTRL-1; EXHIBITS ETHYLENE-
FT MUTAGEN 694 694 TREATED PHENOTYPE.
FT MUTAGEN 694 694 D->E; IN CTRL-1; EXHIBITS ETHYLENE-
SQ SEQUENCE 821 AA; 90306 MW; 29223D3DCDDCC15BC CAC64;
Query Match 34.0%; Score 442; DB 1; Length 821;
Best Local Similarity 40.2%; Pred. No. 3.9e-28;
Matches 99; Conservative 41; Mismatches 90; Indels 16; Gaps 5
QY 1 YAAKNIQDKEVAAYKKLIE-----KEAIIIVLSHRNIIQFYGVILEPNNYGI 50
DB 566 HRAEHWGSG--AAVYILMEQDFHAEVNEFLREVAIMKRLRHPIIVLFWGAVTQPNLSI 623
QY 51 VEEVYASLGSLYDYN-SNRSEEMDMDHMTATDVAKGMYHMEAPVYVTRDLSKRV 109
DB 624 VTEVYSRSGSLYRLILKSGARREQLDERRRLSMAVYDAKGMVYLNENP-PIVARDLSPLV 682
QY 110 VIADGVLIKCDGASRPFNHT--THMSLVGTFPMWABEVISQLPVSECTDYSGVVLV 167
DB 683 LVDKKYTVKVCDFGLSRKASFTPLSSKSAAGPEWMADEVLRDEPSNEKSDVYSGVILM 742
QY 168 EMLTREVPFKLEGLQVAMLVENKNERLTIPSSCSRSAEILLHQWEADAKKRSFQIIT 227
DB 743 EIAITQQEPMGNLNPAAQVAAVGFKKRLRIELPNNLPQVAAIIEGCTWNEPMWRSPFATIM 802
QY 228 SILESM 233
DB 803 DLRLPL 808
RESULT 7
MK37_MOUSE STANDARD; PRT; 579 AA.
AC Q62073;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE M10gen-activated protein kinase kinase kinase 7 (EC 2.7.1.-)
DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-
DE activated kinase 1).
GN MAP3K7 OR TAK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9612327; PubMed=8533096.
RA Yamaguchi K., Shirakabe K., Shibuya H., Irie K., Ohishi I., Ueno N.,
RA Taniuchi T., Nishida E., Matsuno K.;
RT "Identification of a member of the MAPKK family as a potential
RT mediator of TGF-beta signal transduction.";
RL Science 270:2008-2011(1995).
CC -1- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MARKS.
CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B
CC ACTIVATION.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C210RF7.
CC THIS SWISS-PROT entry is copyrighted by its producer through a copyright

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; D76446; BAA1184.1; -  
 DR HSSP; P08631; IAD5.  
 DR MGD; MGI:1346877; Map3K7.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Prodom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00221; STYK; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE\_ST; 1.  
 DR TRANSFERASE; Serine/threonine-protein kinase; ATP-binding.  
 DR DOMAIN 8 16 POLY-SER.  
 DR NP\_BIND 36 291 PROTEIN KINASE.  
 DR BINDING 63 63 ATP (BY SIMILARITY).  
 DR ACT\_SITE 156 156 BY SIMILARITY.  
 DR SEQUENCE 579 AA; 64227 MW; 97C8F6F3C8E283EE CRC64;

Query Match 30.4%; Score 395; DB 1; Length 579;  
 Best Local Similarity 37.1%; Pred. No. 1.6e-24;  
 Matches 91; Conservative 46; Mismatches 80; Indels 28; Gaps 11;

QY 2 RAKWISODKEVAVKLLIKERAE-----ILSVLSHRNIIQFYGVILEPPNGIVT 52  
 DB 52 KAKW--RAKDVAIK---QISESESRKAFIVELRLQLSRVNHPVIYLYGACLNIP--VCLVM 104  
 QY 53 EYASIGSLYDIYNSRSEEM---DMDHITWATDVAKGMHYLHMAAPVKYIHRDLKSRNV 109  
 DB 105 EYAEAGSGLYNYLHIG--AEPLPYTYTAHAMSCLOCSGVAYLHSMQPKALIHRLDKPNNL 162  
 QY 110 VTAADG-VLKICDPG-ASRPFNHTHTMSLVGTFPMMAPEVIQSLPVSETCDTYSYGVVLM 167  
 DB 163 LLVAGGVYLVKICDPGACDIQTHMTNNK--GSAAMMAPEVFEQSGVSEKCDVFSWGIILW 220  
 QY 168 EMLTREVPFKGLEG--LQVAVLVVERKNERLTPSSCPSPFAELLHQWEADAKRPSFKQ 225  
 DB 221 EVITRRKPFDEIGPAPRIWM-AVHNGTRPPLIKNLKPISLMTKRCWSDPSPRSMEE 279

## RESULT 8

M3K7 HUMAN STANDARD; PRT; 606 AA.  
 ID M3K7 HUMAN 043317; 043319;  
 AC 043318; 043317; 043319; Created)  
 DT 16-OCT-2001 (Rel. 40; Last sequence update)  
 DT 15-OCT-2002 (Rel. 41; Last annotation update)  
 DE Mitogen-activated protein kinase kinase 7 (EC 2.7.1.1.)  
 DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-  
 DE activated kinase 1).  
 GN MAP3K7 OR TAK1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RC TISSUE=lung;  
 RC MEDLINE=98153801; PubMed=9480845;  
 RA Sakurai H., Shigemori N., Hasegawa K., Sugita T.;  
 RT "TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an

RT NF-kappa B-inducing kinase-independent mechanism";  
 RL Biochem Biophys. Res Commun. 243:545-549(1998).  
 CC -1- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDERPHOSPHORYLATED  
 CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B  
 CC ACTIVATION.  
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1A, 1B (SHOWN HERE) AND 1C; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MAP KINASE KINASE KINASE SUBFAMILY.  
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AB009357; BAA25026.1; -  
 DR EMBL; AB009356; BAA25025.1; -  
 DR EMBL; AB009358; BAA25027.2; -  
 DR HSSP; P08631; IAD5.  
 DR GeneW; HGNC:6859; MAP3K7.  
 DR MIM; 602614; -  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Prodom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00221; STYK; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE\_ST; 1.  
 DR PROSITE; PS00111; PROTEIN KINASE\_DOM; 1.  
 DR TRANSFERASE; Serine/threonine-protein kinase; ATP-binding;  
 DR KW Alternative splicing.  
 DR FT DOMAIN 8 14 POLY-SER.  
 DR FT NP\_BIND 36 291 PROTEIN KINASE.  
 DR FT BINDING 63 63 ATP (BY SIMILARITY).  
 DR FT ACT\_SITE 156 156 BY SIMILARITY.  
 DR FT VARSPIC 404 430 MISSING (IN ISOFORM 1A).  
 DR FT VARSPIC 509 518 MISSING (IN ISOFORM 1C).  
 DR FT VARSPIC 519 606 MISSING (IN ISOFORM 1C).  
 DR SEQUENCE 606 AA; 67196 MW; 3D8F8147CD174013 CRC64;

Query Match 30.4%; Score 395; DB 1; Length 606;  
 Best Local Similarity 37.1%; Pred. No. 1.7e-24;  
 Matches 91; Conservative 46; Mismatches 80; Indels 28; Gaps 11;

QY 2 RAKWISODKEVAVKLLIKERAE-----ILSVLSHRNIIQFYGVILEPPNGIVT 52  
 DB 52 KAKW--RAKDVAIK---QISESESRKAFIVELRLQLSRVNHPVIYLYGACLNIP--VCLVM 104  
 QY 53 EYASIGSLYDIYNSRSEEM---DMDHITWATDVAKGMHYLHMAAPVKYIHRDLKSRNV 109  
 DB 105 EYAEAGSGLYNYLHIG--AEPLPYTYTAHAMSCLOCSGVAYLHSMQPKALIHRLDKPNNL 162  
 QY 110 VTAADG-VLKICDPG-ASRPFNHTHTMSLVGTFPMMAPEVIQSLPVSETCDTYSYGVVLM 167  
 DB 163 LLVAGGVYLVKICDPGACDIQTHMTNNK--GSAAMMAPEVFEQSGVSEKCDVFSWGIILW 220  
 QY 168 EMLTREVPFKGLEG--LQVAVLVVERKNERLTPSSCPSPFAELLHQWEADAKRPSFKQ 225  
 DB 221 EVITRRKPFDEIGPAPRIWM-AVHNGTRPPLIKNLKPISLMTKRCWSDPSPRSMEE 279

## RESULT 9

ABLI\_HUMAN

ID ABL1 HUMAN STANDARD: PRT: 1130 AA.  
 AC P00519; Q16133; Q13869; Q13870;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Proto-oncogene tyrosine-protein kinase ABL1 (EC 2.7.1.112) (p150)  
 GN (c-abl).  
 GN ABL1 OR ABL OR JTK7.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxId=9606;  
 RN [1]  
 RC TISSUE=FROM N.A.  
 RC TISSUE=FDroblast;  
 RX MEDLINE=90082420; PubMed=2687768;  
 RA Falaschein E., Einat M., Gokkel E., Marcelle C., Croce C.M.,  
 RA Gale R.P., Canaan E.;  
 RT "Nucleotide sequence analysis of human abl and bcr-abl cDNAs";  
 RT Oncogene 4:1477-1481(1989).  
 [2]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RX MEDLINE=87028219; PubMed=3021337;  
 RA Shivelman E., Lifshitz B., Gale R.P., Roe B.A., Canaan E.;  
 RT "Alternative splicing of RNAs transcribed from the human abl gene and  
 RT from the bcr-abl fused gene";  
 RT Cell 47:277-284(1986).  
 [3]  
 RN SEQUENCE FROM N.A. (ISOFORMS IA AND IB).  
 RP TISSUE=Lung carcinoma;  
 RC MEDLINE=95394474; PubMed=7665185;  
 RA Chisese S.L., Bodenteich A., Wang Y.-F., Wang Y.-P., Burtan D.,  
 RA Clifton S.W., Crabtree J., Freeman A., Iyer K., Jhan L., Ma Y.,  
 RA McLaury H.-J., Pan H.-Q., Sarhan O.H., Toth S., Wang Z., Zhang G.,  
 RA Heisterkamp N., Groffen J., Roe B.A.;  
 RT "Sequence and analysis of the human ABL gene, the BCR gene, and  
 RT regions involved in the Philadelphia chromosomal translocation";  
 RT Genomics 27:67-82(1995).  
 [4]  
 RN SEQUENCE OF 360-426 FROM N.A.  
 RP MEDLINE=83245023; PubMed=6191223;  
 RA Groffen J., Heisterkamp N., Reynolds F.H. Jr., Stephenson J.R.;  
 RT "Homology between phosphotyrosine acceptor site of human c-abl and  
 RT viral oncogene products";  
 RT Nature 304:167-169(1983).  
 [5]  
 RP SEQUENCE OF 27-40 FROM N.A.  
 RP MEDLINE=88065859; PubMed=2825022;  
 RA Falaschein E., Marcelle C., Rosner A., Canaan E., Gale R.P.,  
 RA Drazan O., Smith S.D., Croce C.M.;  
 RT "A new fused transcript in Philadelphia chromosome positive acute  
 RT lymphocytic leukemia";  
 RT Nature 330:386-388(1987).  
 [6]  
 RN SEQUENCE OF 825-845 FROM N.A.  
 RP MEDLINE=94142331; PubMed=7545908;  
 RA Inokuchi K., Futaki M., Dan K., Nomura T.;  
 RT "Sequence analysis of the mutation at codon 834 and the sequence  
 RT variation of codon 837 of c-abl gene";  
 RT Leukemia 8:343-344(1994).  
 [7]  
 RN STRUCTURE BY NMR OF SH2 DOMAIN.  
 RP MEDLINE=92370689; PubMed=1505033;  
 RA Overduin M., Rios C.B., Mayer B.J., Baltimore D., Cowburn D.;  
 RT "Three-dimensional solution structure of the src homology 2 domain of  
 RT c-abl";  
 RT Cell 70:697-704(1992).  
 [8]  
 RP STRUCTURE BY NMR OF SH2 DOMAIN.  
 RP MEDLINE=93101586; PubMed=1281542;  
 RA Overduin M., Mayer B.J., Rios C.B., Baltimore D., Cowburn D.;  
 RT "Secondary structure of Src homology 2 domain of c-abl by  
 RT heteronuclear NMR spectroscopy in solution.";

RL Proc. Natl. Acad. Sci. U.S.A. 89:11673-11677(1992).  
 RN [9]  
 RP STRUCTURE BY NMR OF SH3 DOMAIN.  
 RX MEDLINE=96131878; PubMed=8590002;  
 RA Gosser Y.O., Zheng J., Overduin M., Mayer B.J., Cowburn D.;  
 RT "The solution structure of Abl SH3, and its relationship to SH2 in  
 RT the SH(32) construct";  
 RT Structure 3:1075-1086(1995).  
 [10]  
 RN 3D-STRUCTURE MODELING OF SH3 DOMAIN.  
 RX MEDLINE=95199229; PubMed=7892170;  
 RA Pisabarro M.T., Ortiz A.R., Serrano L., Wade R.C.;  
 RT "Homology modeling of the Abl SH3 domain";  
 RT Proteins 20:203-215(1994).  
 [11]  
 RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 57-218.  
 RP MEDLINE=96398698; PubMed=8805596;  
 RA Nam H.-U., Hasegawa W.G., Roberts T.M., Frederick C.A.;  
 RT "Intramolecular interactions of the regulatory domains of the Bcr-Abl  
 RT kinase reveal a novel control mechanism";  
 RT Structure 4:1105-1114(1996).  
 [12]  
 RN X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS) OF 64-121.  
 RP MEDLINE=96365516; PubMed=9698566;  
 RA Pisabarro M.T., Serrano L., Williams M.;  
 RT "Crystal structure of the Abl SH3 domain complexed with a designed  
 RT high-affinity peptide ligand: implications for SH3-ligand  
 RT interactions";  
 RT J. Mol. Biol. 281:513-521(1998).  
 [13]  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; IA (SHOWN HERE) AND IB; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.  
 CC -1- DISEASE: PARTICIPATES IN A T(9;22) (Q34;Q11) CHROMOSOMAL  
 CC TRANSLOCATION THAT PRODUCES A BCR-ABL ONCOGENE RESPONSIBLE FOR  
 CC CHRONIC MYELOID LEUKEMIA (CML), ACUTE MYELOID LEUKEMIA (AML), AND  
 CC ACUTE LYMPHOBLASTIC LEUKEMIA (ALL).  
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL  
 CC SUBFAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC -1- DATABASE: NAME=Atlas Genet. Cyogenet. Oncol. Haematol.;  
 CC WWW="http://www.infobiogen.fr/services/chromocancer/Genes/ABL.html".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X16416; CAA34438.1; -  
 DR EMBL: M14752; AAB51561.1; -  
 DR EMBL: U07563; AAB60393.1; -  
 DR EMBL: U07563; AAB60393.1; -  
 DR EMBL: U07561; AAB60393.1; JOINED.  
 DR EMBL: S69223; AAB14034.1; -  
 DR EMBL: A25582; TVHDA.  
 DR PDB: 1AB2; 31-JAN-94.  
 DR PDB: 1AB1; 01-NOV-94.  
 DR PDB: 2AB1; 04-SEP-97.  
 DR PDB: 1AWO; 28-JAN-98.  
 DR PDB: 1BB2; 25-NOV-98.  
 DR GeneW; HGNC:76; ABL1.  
 DR MIM; 189980; -  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00017; SH2; 1.









```

Db 419 ENHVKVADFGSLMTGDTYTAHAGAKFPKMTAPESLAVTFSTKSVMAFGVLLWEI 478
Qy 170 LTRREV-PFKGLEGLQVAMLVVEKNERLTTPSSCPRSPFAELLHQCEADAKKPSFKQIIS 228
Db 479 ATYGMSPYRGIDLSQV-YDLLEKGYRMGEPCGCPKPVYELMAACKWSPADPSPFAETHQ 537
Qy 229 ILESMSNDTSLPDK 242
Db 538 AFETMFHDSISEE 551

RESULT 13
ABL_DROME STANDARD; PRT; 1520 AA.
ID ABL_DROME
AC P00522;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DT Tyrosine-protein kinase Abl (EC 2.7.1.112) (D-ash).
ABL OR DASH OR ABL-1.
Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
NCBI_Taxid=7227;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=86174726; PubMed=2832740;
RA Henikoff M.J., Bennett R.L., Gertler F.B., Hoffman F.M.;
RT "DNA sequence, structure, and tyrosine kinase activity of the
RT Drosophila melanogaster Abelson proto-oncogene homolog.";
RL Mol. Cell. Biol. 8:843-853(1988).
[2]
RN SEQUENCE OF 374-648 FROM N.A.
RX MEDLINE=84082064; PubMed=6317185;
RA Hoffmann F.M., Freese L.D., Hoffman-Palk H., Shilo B.-Z.;
RT "Nucleotide sequences of the Drosophila src and abl homologs:
RT conservation and variability in the src family oncogenes.";
RL Cell 35:393-401(1983).
[3]
RN FUNCTION.
RX MEDLINE=98298928; PubMed=9635189;
RA Loureiro J., Pellet M.;
RT "Role of Armadillo, a Drosophila catenin, during central nervous
RT system development.";
RL Curr. Biol. 8:622-632(1998).
-1- FUNCTION: ARM AND ABL PROTEINS FUNCTION COOPERATIVELY AT ADHERENS
JUNCTIONS IN BOTH THE CNS AND EPIDERMIS.
-1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
tyrosine phosphate.
-1- SUBCELLULAR LOCATION: Cytoplasmic.
-1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL
SUBFAMILY.
-1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
-1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
```

```

DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR ProDom; PD000066; SH3; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PSS0001; SH2; 1.
DR PROSITE; PSS0002; SH3; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW SH2 domain; SH3 domain.
FT DOMAIN 204 265 SH3.
FT DOMAIN 271 363 SH2.
FT DOMAIN 388 644 PROTEIN KINASE.
FT NP_BIND 394 402 ATP (BY SIMILARITY).
FT BINDING 417 417 ATP (BY SIMILARITY).
FT ACT_SITE 509 509 BY SIMILARITY.
FT MOD_RES 539 539 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CONFLICT 374 377 LSPK -> ASAQ (IN REF. 2).
FT CONFLICT 645 648 ESS1 -> VGDV (IN REF. 2).
SQ SEQUENCE 1520 AA; 161836 MW; AD6A5060579PAD7B CRC64;

Query Match 28.0%; Score 364.5; DB 1; Length 1520;
Best local Similarity 34.1%; Pred. No. 1.4e-21;
Matches 86; Conservative 47; Mismatches 104; Indels 15; Gaps 7;

Qy 1 YRAKMSIQDEKVAVKL-----LK-IEKAEILSVLSHRNIIOFVGVILEPPNYGVITE 53
Db 403 YEAVWKRYGVNTAVKTLKEDPTMLKDFLEBAIMKEKHNLVQLIGVCRRPPFIITE 462
Qy 54 YASIGSLYDYNRSSEEMDMHTATDVAKGMHYLMEAPKVIHRDLKSRNVIAA 113
Db 463 FMSHGNLDPLRSAGRETLDAVALLVYATQIASGMSYLESR---NYIHRDLAARNCLVGD 519
Qy 114 DGVLKICDFGASRPHNTHTMSLVGT-FP--WMAPEVIOQLPVSCTDPTYSYGVVLEML 170
Db 520 NKLKVVADFGSLARMDDTYTAHAGAKFPKMTAPESLAVNKFSTYSVVAFGVLLWEIA 579
Qy 171 TREV-PFKGLEGLQVAMLVVEKNERLTTPSSCPRSPFAELLHQCEADAKKPSFKQIISI 229
Db 580 TYGMSPYRAID-LTDVYHKLDKGYRMERPCGCPPEVYDLMRCQWMDATDRPFRKSIHHA 638
Qy 230 ILESMSNDTSLPD 241
Db 639 LEHMFQESSITE 650

RESULT 14
ID KYKI_DICDI STANDARD; PRT; 1584 AA.
AC KYKI_DICDI
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Non-receptor tyrosine kinase spore lysin A (EC 2.7.1.112) (Tyrosine-
DE protein kinase 1).
GN PYKA OR SPKA OR DEPK1.
OC Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_Taxid=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JH10;
```

RX MEDLINE=97053827; PubMed=8898241; Spudich J.A.;  
 RA Nickolls G.H.; Oshero N.; Loomis W.F.; Spudich J.A.;  
 RT "The Dictyostelium dual-specificity kinase *spkA* is essential for  
 RT spore differentiation.";  
 RL Development 122:3295-3305 (1996).  
 RN [2]  
 RP SEQUENCE OF 1248-1584 FROM N.A.  
 RX MEDLINE=90287147; PubMed=1972546;  
 RA Tan J.L.; Spudich J.A.;  
 RT "Developmentally regulated protein-tyrosine kinase genes in  
 RT dictyostelium discoideum.";  
 RL Mol. Cell. Biol. 10:3578-3583 (1990).  
 CC -1- FUNCTION: ESSENTIAL FOR SPORE DIFFERENTIATION.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT WITH A PEAK  
 CC DURING THE MOUND STAGE OF MORPHOGENESIS.  
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.  
 CC -1- SIMILARITY: CONTAINS 1 SAM DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U32174; AAB41125.1; -;  
 DR EMBL; M33785; AAB33202.1; -;  
 DR PIR; A35670; A35670.  
 DR Dictydb; DD03010; pykA.  
 DR InterPro; IPR000719; Euk\_Pkinase.  
 DR InterPro; IPR001660; SAM.  
 DR InterPro; IPR003878; SPRY domain.  
 DR InterPro; IPR003877; SPRY receptor.  
 DR InterPro; IPR004040; STY\_Pkinase.  
 DR InterPro; IPR001245; Tyr\_Pkinase.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR Pfam; PF00536; SAM; 1.  
 DR Pfam; PF00622; SPRY; 3.  
 DR ProDom; PD000001; Euk\_Pkinase; 1.  
 DR SMART; SMO0454; SAM; 1.  
 DR SMART; SMO0449; SPRY; 1.  
 DR SMART; SMO0221; STYK; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS50109; PROTEIN KINASE TYR; 1.  
 DR PROSITE; PS50105; SAM DOMAIN; 1.  
 RT Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.  
 FT DOMAIN 908 972 SAM.  
 FT 403 420 POLY-ASN.  
 FT 428 435 POLY-THR.  
 FT 449 480 POLY-ASN.  
 FT 483 491 POLY-ASN.  
 FT 494 508 POLY-ASN.  
 FT 512 532 POLY-ASN.  
 FT 536 600 POLY-ASN.  
 FT 608 811 POLY-PHE.  
 FT 1026 1029 POLY-SER.  
 FT 1195 1210 POLY-ASN.  
 FT 1215 1220 POLY-GLN.  
 FT 1224 1233 POLY-GLN.  
 FT 1266 1274 POLY-PRO.  
 FT 1289 1561 PROTEIN KINASE.  
 FT 1295 1303 ATP (BY SIMILARITY).  
 FT 1316 1316 ATP (BY SIMILARITY).  
 FT 1417 1417 BY SIMILARITY.  
 FT 1428 1428 D -> R (IN REF. 2).  
 FT 1435 1435 V -> L (IN REF. 2).  
 FT 1584 AA; 174304 MW; 5D1589458DBE01E3 CRC64;  
 SQ SEQUENCE

Query Match 27.3%; Score 355.5; DB 1; Length 1584;

Best Local Similarity 34.1%; Pred. No. 7,8e-21;  
 Matches 93; Conservative 47; Mismatches 96; Indels 37; Gaps 8;  
 QY 2 RAKWISQDKEVAVK-----KLTKIKGAELTSLVSHRNITQFGVILE--PPNY 48  
 DB 1305 RGYW--RETQVAIKIIRYDQFKTKSLVMFQNEVGLSKLRHNVQFQAGCTAGEDIH 1362  
 QY 49 GIYTEXASLGSLDYDIYNSNRSEEMDMHI-MTWTADVAKGMHYLHAEAVKVIHRDLKSR 107  
 DB 1363 CITEWNGGSLQFLTDHNLLEQNPRIKLADIAKGMVTHGWT-PILHRDLSR 1421  
 QY 108 NVYI-----AAGVLKICDPGASRPNNHTTM--SLVTFPMAPAEVYQSLP 152  
 DB 1422 NILLDHNIDKPNVVSRODIKCKISDFGLSRKKEQASQMTQSGCIPYMAPEVFKGDS 1481  
 QY 153 VSETCTYISGVVLNEMLTREVPFKLEGQVAMLVENKNERITPSCPSFAELHOC 212  
 DB 1482 NSEKSDVSYGMVLFELTSDPEQDMKPKMHLAYESYRPIPLTSSKMKELTQC 1541  
 QY 213 WEADAKRPSFKQIISLBSM-----SNDTSLP 240  
 DB 1542 WDSNPDSRPFTFKQIIVHLKEMEDQGVSSPASVP 1574  
 RESULT 15  
 ID ABL\_PSVHY STANDARD; PRT; 439 AA.  
 AC P10447;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tyrosine-protein kinase transforming protein ABL (BC 2.7.1.112).  
 GN V-ABL.  
 OS Feline sarcoma virus (strain Hardy-Zuckerman 2).  
 OC Viruses; Retroviral viruses; Retroviridae; Gammaretrovirus.  
 OC NCBI\_TaxID=11776;  
 RX MEDLINE=67141338; PubMed=3029415;  
 RA Bergold P.J.; Blumenthal J.A.; D'Andrea E.; Snyder H.W. Jr.;  
 RA Lederman L.; Silverstone A.; Nguyen H.; Besmer P.;  
 RT "Nucleic acid sequence and oncogenic properties of the H22 feline  
 RT sarcoma virus v-*abl* insert.";  
 RL J. Virol. 61:1193-1202 (1987).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-ABL-POL  
 CC POLYPROTEIN.  
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ABL  
 CC SUBFAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M15805; AAA43042.1; -;  
 DR PIR; A26132; TVWVAB.  
 DR HSP; P00519; 1BBZ.  
 DR InterPro; IPR000719; Euk\_Pkinase.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001245; Tyr\_Pkinase.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR Pfam; PF00401; SH2DOMAIN.  
 DR PRINTS; PR00109; TYRKINASE.

DR ProDom; PD000001; Euk\_pkinase; 1.  
DR ProDom; PD000066; SH3; 1.  
DR ProDom; PD000093; SH2; 1.  
DR SMART; SM00252; SH2; 1.  
DR SMART; SM00326; SH3; 1.  
DR SMART; SM00219; TYRC; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS50001; SH2; 1.  
DR PROSITE; PS50002; SH3; 1.  
KW Polyprotein; Tyrosine-protein kinase; Transferase; Oncogene;  
KW SH2 domain; SH3 domain.  
FT DOMAIN 10 70 SH3.  
FT DOMAIN 76 166 SH2.  
FT DOMAIN 191 439 PROTEIN\_KINASE.  
SQ SEQUENCE 439 AA; 50004 MW; 13579EDFED1481AB CRC64;

Query Match 27.0%; Score 351.5; DB 1; Length 439;  
Best Local Similarity 32.9%; Pred. No. 3.6e-21;  
Matches 78; Conservative 51; Mismatches 93; Indels 15; Gaps 7;

QY 1 YRAKWIISODPEVAVVKL---LKIE---KEAEILSVLSHRNIIQFYGVILEPPNYGIIVTE 53  
DB 206 YEGVWKXKSLTVAVVKLKEDETVMEVEEPLKEAAYMKEIKGNILVQLGVCTRREPPYITTE 265  
QY 54 YASLSGLVDYINSNRSEKMDHIMTWATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAA 113  
DB 266 FWTYGNLDYLRCCNRQOEYNAVLLYMATQIISAMEYLEK--NFIHRDLAARNCVIGE 322  
QY 114 DGLKICPEGASRFNHTTHMSLVGT-PP--WMAPEVIQSLPVSETCCTYSYGVVLEML 170  
DB 323 NHIVVAVADFGLSRLMTGDIYTAHAGTKPPKWTAPESLAYNKFISIKSDVMAFGVILMEIA 382  
QY 171 TRGV-PFNGLEGLQVAVLVEKNERLTIIPSCPRSFALHQCWEADAKKRPSPFOI 226  
DB 383 TYGMSFPYFGIDLSQY-YELLEKDYMERPEGCEKVEYELMRAQWQMNPSDRPAFAEI 438

Search completed: May 1, 2003, 20:36:20  
Job time : 15.4017 secs.



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:28:47 ; Search time 27.5593 Seconds

(without alignments)  
1846.697 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_31\_277

Perfect score: 1300

Sequence: 1 TRAKISQDKEVAVKLIK.....SILESMSNDTSLPDKNSFL 247

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP\_ARCHAEA:\*  
2: SP\_BACTERIA:\*  
3: SP\_FUNGI:\*  
4: SP\_HUMAN:\*  
5: SP\_INVERTEBRATE:\*  
6: SP\_MAMMAL:\*  
7: SP\_MHC:\*  
8: SP\_ORGANELLE:\*  
9: SP\_PHAGE:\*  
10: SP\_PLANT:\*  
11: SP RODENT:\*  
12: SP\_VIRUS:\*  
13: SP\_VERTEBRATE:\*  
14: SP\_UNCLASSIFIED:\*  
15: SP\_VIRUS:\*  
16: SP\_BACTERIAP:\*  
17: SP\_ARCHAEP:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1300	100.0	455	4 Q9HCC4	Q9HCC4 homo sapien
2	1300	100.0	800	4 Q9HDD2	Q9HDD2 homo sapien
3	1300	100.0	800	4 Q9HCC5	Q9HCC5 homo sapien
4	1300	100.0	800	4 Q9NYL2	Q9NYL2 homo sapien
5	1300	100.0	800	4 Q9NYE9	Q9NYE9 homo sapien
6	1289	99.2	454	11 Q9ESL3	Q9ESL3 mus musculu
7	1289	99.2	802	11 Q9ESL4	Q9ESL4 mus musculu
8	1064	81.8	371	13 Q902Y8	Q902Y8 brachydanio
9	511.5	39.3	406	10 Q8S9K4	Q8S9K4 arabidopsis
10	511.5	39.3	406	10 Q8S9K4	Q8S9K4 arabidopsis
11	508.5	39.1	1066	4 Q9H2N5	Q9H2N5 homo sapien
12	495.5	38.1	740	5 Q21982	Q21982 caenorhabd
13	494	38.0	1001	11 Q8VDG6	Q8VDG6 mus musculu
14	490	37.7	859	4 Q8W125	Q8W125 homo sapien
15	483	37.2	564	4 Q9H1Y7	Q9H1Y7 homo sapien
16	482.5	37.1	847	4 Q16584	Q16584 homo sapien

17	481.5	37.0	850	11 Q9J115	Q9J115 mus musculu
18	480	36.9	886	10 Q9Y1I8	Q9Y1I8 arabidopsis
19	475.5	36.6	1030	10 Q9C9U5	Q9C9U5 arabidopsis
20	474	36.5	570	4 Q8WMN2	Q8WMN2 homo sapien
21	474	36.5	1036	4 Q8WMN1	Q8WMN1 homo sapien
22	472.5	36.3	977	5 Q9VW24	Q9VW24 drosoophila
23	469.5	36.1	966	4 Q43283	Q43283 homo sapien
24	460.5	35.4	462	10 Q39886	Q39886 glycine max
25	453	34.8	982	10 Q65833	Q65833 lycopersico
26	452	34.8	903	10 Q9FPR5	Q9FPR5 oryza sativ
27	448	34.5	412	10 Q9M085	Q9M085 arabidopsis
28	445.5	34.3	1020	5 Q9M313	Q9M313 drosoophila
29	445.5	34.3	1148	5 Q95VF6	Q95VF6 drosoophila
30	445.5	34.3	1161	5 Q95UN8	Q95UN8 drosoophila
31	441	33.9	525	10 Q9P1L6	Q9P1L6 arabidopsis
32	440	33.8	962	10 Q93Y88	Q93Y88 lycopersico
33	438	33.7	525	10 Q8W022	Q8W022 arabidopsis
34	436.5	33.6	933	10 Q9FPR3	Q9FPR3 arabidopsis
35	435.5	33.5	847	10 Q93XL9	Q93XL9 rosa hybrid
36	434	33.4	855	5 Q01700	Q01700 caenorhabd
37	432	33.2	957	10 Q9FPR4	Q9FPR4 hordeum vul
38	431.5	33.2	523	10 Q81808	Q81808 arabidopsis
39	431.5	33.2	570	10 Q8RML6	Q8RML6 arabidopsis
40	431	33.2	829	10 Q24027	Q24027 lycopersico
41	427	32.8	475	10 Q9STG5	Q9STG5 arabidopsis
42	427	32.8	637	10 Q94J41	Q94J41 oryza sativ
43	426	32.8	480	10 Q9PDV7	Q9PDV7 laqus sylva
44	422.5	32.5	2651	10 Q9FRR5	Q9FRR5 arabidopsis
45	421	32.4	806	10 Q9ZSD8	Q9ZSD8 lycopersico

## ALIGNMENTS

RESULT 1  
Q9HCC4 PRELIMINARY; PRT; 455 AA.  
ID Q9HCC4  
AC Q9HCC4;  
DT 01-MAR-2001 (T-EMBLrel. 16, Created)  
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)  
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
DE MTRK-beta (Similar to sterile-alpha motif and leucine zipper  
DE containing kinase AZK) (Mixed lineage kinase) (Mixed lineage  
DE kinase-related kinase MTRK-beta).  
GN MTRK.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gotch I., Adachi M., Nishida E.;  
RT "Identification and Characterization of a Novel MAP Kinase Kinase  
RT kinase, MTRK.";  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA TISSUE-COLON;  
RA Strausberg R.;  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Acton S.;  
RT "MLK-mixed lineage kinase.";  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=21950776; PubMed=11836244;  
RA Gross E.A., Callow M.G., Waldbaum L., Thomas S., Ruggieri R.;  
RT "MRK, a Mixed Lineage Kinase-Related Molecule That Plays a Role in  
RT gamma-Radiation-Induced Cell Cycle Arrest.";  
RL J. Biol. Chem. 277:13873-13882(2002).  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL: AB049734; BAB16445.1; -  
 DR EMBL: BC001401; AAH01401.1; -  
 DR EMBL: AF325454; AAK11615.1; -  
 DR EMBL: AF480462; AAL85892.1; -  
 DR HSSP: P12931; 1PMK.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR004040; STY\_pkinase.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00069; pkinase.1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR PRODOM: PD000001; Euk\_pkinase.1.  
 DR SMART: SM00221; STYK; 1.  
 DR SMART: SM00220; S\_TKC; 1.  
 DR SMART: SM00219; TYRK; 1.  
 DR PROSITE: PS0011; PROTEIN KINASE DOM; 1.  
 DR PROSITE: PS00108; PROTEIN KINASE ST; 1.  
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 455 AA; 51582 MW; B87D84A4D58B752 CRC64;

Query Match 100.0%; Score 1300; DB 4; Length 455;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-114;  
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWI SODKEVAVKKLKI EKEAEILSVLSHRNIIQFYGVILEPPNNGIVTEYASLSGL 60  
 DB 31 YRAKWI SODKEVAVKKLKI EKEAEILSVLSHRNIIQFYGVILEPPNNGIVTEYASLSGL 90  
 QY 61 YDYNINRSEEMDMHMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVIAADGVLKIC 120  
 DB 91 YDYNINRSEEMDMHMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVIAADGVLKIC 150  
 QY 121 DFGASRFHNHTTMSLVGTFFPMWAPVIOQLPVSECTDITYSGVILWMLTREVPFKGLE 180  
 DB 151 DFGASRFHNHTTMSLVGTFFPMWAPVIOQLPVSECTDITYSGVILWMLTREVPFKGLE 210  
 QY 181 GLOVAMLVKERNRLTIPSSCPSPFAELHQCWEADAKRPSFKOIIISLESMSNDTSLP 240  
 DB 211 GLOVAMLVKERNRLTIPSSCPSPFAELHQCWEADAKRPSFKOIIISLESMSNDTSLP 270  
 QY 241 DKCNSFL 247  
 DB 271 DKCNSFL 277

RESULT 2  
 Q9HDD2 PRELIMINARY; PRT; 800 AA.  
 ID Q9HDD2  
 AC 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Placible mixed-lineage kinase protein.  
 GN MLKXK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCB1\_TaxID=9606;  
 RX SEQUENCE FROM N.A.  
 RP TISSUE=LIMPHOID ORGAN;  
 RC ABE Y., Ueda N.;  
 RA "Placible Mixed-lineage kinase derived from LAK cell."  
 RT Submitted (JUL-1999) to the EMBL/GenBank/DBD databases.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL: AB030034; BAB12040.1; -  
 DR HSSP: P12931; 1PMK.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR001660; SAM.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR004040; STY\_pkinase.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00069; pkinase.1.

DR Pfam: PF000536; SAM; 1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR PRODOM: PD000001; Euk\_pkinase.1.  
 DR SMART: SM00454; SAM; 1.  
 DR SMART: SM00221; STYK; 1.  
 DR SMART: SM00220; S\_TKC; 1.  
 DR SMART: SM00219; TYRK; 1.  
 DR PROSITE: PS0011; PROTEIN KINASE DOM; 1.  
 DR PROSITE: PS00108; PROTEIN KINASE ST; 1.  
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 800 AA; 91155 MW; B2814509E54B07A CRC64;

Query Match 100.0%; Score 1300; DB 4; Length 800;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-114;  
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWI SODKEVAVKKLKI EKEAEILSVLSHRNIIQFYGVILEPPNNGIVTEYASLSGL 60  
 DB 31 YRAKWI SODKEVAVKKLKI EKEAEILSVLSHRNIIQFYGVILEPPNNGIVTEYASLSGL 90  
 QY 61 YDYNINRSEEMDMHMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVIAADGVLKIC 120  
 DB 91 YDYNINRSEEMDMHMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVIAADGVLKIC 150  
 QY 121 DFGASRFHNHTTMSLVGTFFPMWAPVIOQLPVSECTDITYSGVILWMLTREVPFKGLE 180  
 DB 151 DFGASRFHNHTTMSLVGTFFPMWAPVIOQLPVSECTDITYSGVILWMLTREVPFKGLE 210  
 QY 181 GLOVAMLVKERNRLTIPSSCPSPFAELHQCWEADAKRPSFKOIIISLESMSNDTSLP 240  
 DB 211 GLOVAMLVKERNRLTIPSSCPSPFAELHQCWEADAKRPSFKOIIISLESMSNDTSLP 270  
 QY 241 DKCNSFL 247  
 DB 271 DKCNSFL 277

RESULT 3  
 Q9HCC5 PRELIMINARY; PRT; 800 AA.  
 ID Q9HCC5  
 AC 09HCC5;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE MLTK-alpha.  
 GN MLTK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCB1\_TaxID=9606;  
 RX SEQUENCE FROM N.A.  
 RP MEDLINE=2164927; Pubmed=11042189;  
 RA Gotoh I., Adachi M., Nishida E.;  
 RT "Identification and Characterization of a Novel MAP Kinase Kinase  
 Kinase, MLTK.";  
 RT J. Biol. Chem. 276:4276-4286(2001).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL: AB049733; BAB16444.1; -  
 DR HSSP: P12931; 1PMK.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR001660; SAM.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR004040; STY\_pkinase.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00069; pkinase.1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR PRODOM: PD000001; Euk\_pkinase.1.  
 DR SMART: SM00454; SAM; 1.  
 DR SMART: SM00221; STYK; 1.  
 DR SMART: SM00220; S\_TKC; 1.  
 DR SMART: SM00219; TYRK; 1.



DR PROSITE; PSS0011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PSS0108; PROTEIN KINASE ST; 1.  
 KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 800 AA; 91188 MW; 2C8593824AB3FADD CRC64;

Query Match 100.0%; Score 1300; DB 4; Length 800;  
 Best Local Similarity 100.0%; Pred. No. 3, 1e-114;  
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKMSIQDEVAVKLLKIKKEKELISVLSHRNIIOFYGVILEPPNYGIVTEYASIGSL 60  
 DB 31 YRAKMSIQDEVAVKLLKIKKEKELISVLSHRNIIOFYGVILEPPNYGIVTEYASIGSL 90  
 QY 61 YDIYNSNSEEMDMHMTATDVAKGMHYLHMEAPVKVIRHDLKSRNVVIAADGVLTIC 120  
 DB 91 YDIYNSNSEEMDMHMTATDVAKGMHYLHMEAPVKVIRHDLKSRNVVIAADGVLTIC 150  
 QY 121 DFGASRFHNHTTHMSLVGTFPPMAPEVIOSLPVSETCDITYSYGVLMEMLTREVPFGLE 180  
 DB 151 DFGASRFHNHTTHMSLVGTFPPMAPEVIOSLPVSETCDITYSYGVLMEMLTREVPFGLE 210  
 QY 181 GIOVAMLVENKNERLTIPSSCPSPFAELLHQCWEADAKKRPFKQIISLESMSNDTSLP 240  
 DB 211 GIOVAMLVENKNERLTIPSSCPSPFAELLHQCWEADAKKRPFKQIISLESMSNDTSLP 270  
 QY 241 DKCNSFL 247  
 DB 271 DKCNSFL 277

## RESULT 4

Q9NYL2 PRELIMINARY; PRT; 800 AA.

AC Q9NYL2; 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Mixed lineage kinase.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;

RP SEQUENCE FROM N.A.  
 RX MEDLINE=20384179; PubMed=10924358;  
 RA Liu T.C., Huang C.J., Chu Y.C., Wei C.C., Chou C.C., Chou M.Y.,  
 RA Chou C.K., Yang J.U.;  
 RT "Cloning and expression of ZAK, a mixed lineage kinase-like protein  
 containing a leucine-zipper and a sterile-alpha motif.";  
 RT Blochem. Biophys. Res. Commun. 274:811-816(2000).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL; AF238255; AAF63490.1; -;  
 DR HSSP; P12931; 1FMK.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001660; SAM.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR Pfam; PF00536; SAM; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00454; SAM; 1.  
 DR SMART; SM00221; STYKC; 1.  
 DR PROSITE; PSS0011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PSS0108; PROTEIN KINASE ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 800 AA; 91264 MW; DA82D7ABB2082F43 CRC64;

Query Match 100.0%; Score 1300; DB 4; Length 800;  
 Best Local Similarity 100.0%; Pred. No. 3, 1e-114;  
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKMSIQDEVAVKLLKIKKEKELISVLSHRNIIOFYGVILEPPNYGIVTEYASIGSL 60  
 DB 31 YRAKMSIQDEVAVKLLKIKKEKELISVLSHRNIIOFYGVILEPPNYGIVTEYASIGSL 90

QY 61 YDIYNSNSEEMDMHMTATDVAKGMHYLHMEAPVKVIRHDLKSRNVVIAADGVLTIC 120  
 DB 91 YDIYNSNSEEMDMHMTATDVAKGMHYLHMEAPVKVIRHDLKSRNVVIAADGVLTIC 150  
 QY 121 DFGASRFHNHTTHMSLVGTFPPMAPEVIOSLPVSETCDITYSYGVLMEMLTREVPFGLE 180  
 DB 151 DFGASRFHNHTTHMSLVGTFPPMAPEVIOSLPVSETCDITYSYGVLMEMLTREVPFGLE 210  
 QY 181 GIOVAMLVENKNERLTIPSSCPSPFAELLHQCWEADAKKRPFKQIISLESMSNDTSLP 240  
 DB 211 GIOVAMLVENKNERLTIPSSCPSPFAELLHQCWEADAKKRPFKQIISLESMSNDTSLP 270  
 QY 241 DKCNSFL 247  
 DB 271 DKCNSFL 277

## RESULT 5

Q9NYE9 PRELIMINARY; PRT; 800 AA.

AC Q9NYE9; 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Sterile-alpha motif and leucine zipper containing kinase AZK (Mixed  
 lineage kinase-related kinase MRK-alpha).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;

RP SEQUENCE FROM N.A.  
 RA McNea J.J., Frima N., Diamond T.E., Dower S.K., Guesdon F.;  
 RT "Cloning and characterization of AZK, a mixed lineage kinase  
 containing a sterile-alpha motif.";  
 RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RL [2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=21950776; PubMed=11836244;  
 RA Gross E.A., Callow M.G., Waldbaum L., Thomas S., Ruggieri R.;  
 RT "MRK, a Mixed Lineage Kinase-related Molecule That Plays a Role in  
 gamma-Radiation-Induced Cell Cycle Arrest.";  
 RT J. Biol. Chem. 277:13873-13882(2002).

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AF251441; AAF65822.1; -;  
 DR EMBL; AF480461; AAL85891.1; -;  
 DR HSSP; P12931; 1FMK.

DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001660; SAM.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR Pfam; PF00536; SAM; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00454; SAM; 1.  
 DR SMART; SM00221; STYKC; 1.  
 DR PROSITE; PSS0011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PSS0108; PROTEIN KINASE ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 800 AA; 9181 MW; B289D835EC52E25 CRC64;

Query Match 100.0%; Score 1300; DB 4; Length 800;  
 Best Local Similarity 100.0%; Pred. No. 3, 1e-114;  
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db 31 YRAKWIQDKEVAVKLLKIEKEAIIISVLSHRNIIQFYGVILLEPPNYGIVTEVASIGSL 90
QY 61 YDYINSNRSEMDMDHMTATADYAKGMHYLHMEAPYKVHRLDKSRNVVIAADGVKIC 120
Db 91 YDYINSNRSEMDMDHMTATADYAKGMHYLHMEAPYKVHRLDKSRNVVIAADGVKIC 150
QY 121 DFGASRFNHTTHMSLVGTFFPMAPEVIOSLPVSETCDTYSYGVVLEMLTREVPFKGLE 180
Db 151 DFGASRFNHTTHMSLVGTFFPMAPEVIOSLPVSETCDTYSYGVVLEMLTREVPFKGLE 210
QY 181 GLOVAMLVVEKNERLITPSSCPSPFAELHOCWEADAKKRSFKOIIISLESMSNDTSLP 240
Db 211 GLOVAMLVVEKNERLITPSSCPSPFAELHOCWEADAKKRSFKOIIISLESMSNDTSLP 270
QY 241 DKNSFL 247
Db 271 DKNSFL 277

```

## RESULT 6

```

ID Q9ESL3 PRELIMINARY; PRT; 454 AA.
AC Q9ESL3;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE MLTK-beta.
GN ZAK OR MLTK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21264927; PubMed=11042189;
RA Gotoh I., Adachi M., Nishida E.;
RT "Identification and Characterization of a Novel MAP Kinase Kinase
RT Kinase, MLTK."
RL J. Biol. Chem. 276:4276-4286(2001).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB049732; BAB16443.1; -.
DR HSSP; P12931; 1FMK.
DR MGD; MGI:1931274; ZAK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKc; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 454 AA; 51366 MW; 35CFCOD729D9395 CRC64;

```

```

Query Match 99.2%; Score 1289; DB 11; Length 454;
Best Local Similarity 98.8%; Pred. No. 1,6e-113;
Matches 244; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 181 GLOVAMLVVEKNERLITPSSCPSPFAELHOCWEADAKKRSFKOIIISLESMSNDTSLP 240
Db 211 GLOVAMLVVEKNERLITPSSCPSPFAELHOCWEADAKKRSFKOIIISLESMSNDTSLP 270
QY 241 DKNSFL 247
Db 271 DKNSFL 277

```

## RESULT 7

```

ID Q9ESL4 PRELIMINARY; PRT; 802 AA.
AC Q9ESL4;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE MLTK alpha.
GN ZAK OR MLTK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21264927; PubMed=11042189;
RA Gotoh I., Adachi M., Nishida E.;
RT "Identification and Characterization of a Novel MAP Kinase Kinase
RT Kinase, MLTK."
RL J. Biol. Chem. 276:4276-4286(2001).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB049731; BAB16442.1; -.
DR HSSP; P12931; 1FMK.
DR MGD; MGI:1931274; ZAK.
DR InterPro; IPR000194; Arpase_a/bcentre.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001650; SAM.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Euk_pkinase; 1.
DR SMART; SM00454; SAM; 1.
DR SMART; SM00221; STYKc; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS00152; ARPASE ALPHA BETA; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 802 AA; 91719 MW; D43IDF8F312A43C CRC64;

```

```

Query Match 99.2%; Score 1289; DB 11; Length 802;
Best Local Similarity 98.8%; Pred. No. 3,4e-113;
Matches 244; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 YRAKWIQDKEVAVKLLKIEKEAIIISVLSHRNIIQFYGVILLEPPNYGIVTEVASIGSL 60
Db 31 YRAKWIQDKEVAVKLLKIEKEAIIISVLSHRNIIQFYGVILLEPPNYGIVTEVASIGSL 90
QY 61 YDYINSNRSEMDMDHMTATADYAKGMHYLHMEAPYKVHRLDKSRNVVIAADGVKIC 120
Db 91 YDYINSNRSEMDMDHMTATADYAKGMHYLHMEAPYKVHRLDKSRNVVIAADGVKIC 150
QY 121 DFGASRFNHTTHMSLVGTFFPMAPEVIOSLPVSETCDTYSYGVVLEMLTREVPFKGLE 180
Db 151 DFGASRFNHTTHMSLVGTFFPMAPEVIOSLPVSETCDTYSYGVVLEMLTREVPFKGLE 210
QY 181 GLOVAMLVVEKNERLITPSSCPSPFAELHOCWEADAKKRSFKOIIISLESMSNDTSLP 240
Db 211 GLOVAMLVVEKNERLITPSSCPSPFAELHOCWEADAKKRSFKOIIISLESMSNDTSLP 270

```

QY 241 DKNSFL 247  
DB 271 DQNSFL 277

## RESULT 8

ID Q902Y8 PRELIMINARY; PRT; 371 AA.  
AC Q902Y8;  
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)  
DE Protein kinase Npk.  
OS Brachydanio rerio (Zebrafish) (Zebra danio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
NCBI\_TaxId=7955;  
[1]

## SEQUENCE FROM N.A.

Chou C.-M., Lee I.-L., Leu J.-H., Huang C.-D.;  
"A novel protein kinase, ZPK, from the zebrafish";  
Submitted (MAY-2000) to the EMBL/Genbank/DBS databases.  
DR EMBL; AF253343; AK52416.1; -  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR PROSITE; PSS0011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PSS0108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
KM ATP-binding; Kinase; Transferase.  
SQ SEQUENCE 371 AA; 42456 MW; 9B918B8AB20D296 CRC64;

Query Match 81.8%; Score 1064; DB 13; Length 371;  
Best Local Similarity 79.8%; Pred. No. 2.4e-92;  
Matches 197; Conservative 23; Mismatches 27; Indels 0; Gaps 0;

QY 1 YAAKVISQKEVAVVKKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYGIVTEVASLQSL 60  
DB 62 YAAHVAVPODEYAVVKKLLKIDAEAILSVLSHKNIIOFGALLLEAPNGIVTEVASRQSL 121  
QY 61 YDIYNSNSEMDMHIMTATDVAKGMHYLHMEAPVYVYIHRDLKSRNVVLAADGVKIC 120  
DB 122 YEYLSASSEEDMDQVMTWAMEIKGMHYLAEPVYVYIHRDLKSRNVVLAADGVKIC 181  
QY 121 DFGASRFNHTTHMSLVGTFFPMABEVIQSLPVSECTDYSYGVVLAEMLTREVFPKGL 180  
DB 182 DFGASKVSHHTHMSLVGTFFPMABEVIQSLPVSECTDYSYGVVLAEMLTREVFPKGL 241  
QY 181 GLOVAVLVEKNERLTITSSCPSPFAELLHQWEADAKRPSFKQIISLESMSNDTSLP 240  
DB 242 GLOVAVLVEKNERLTITSSCPSPFAELLHQWEADAKRPSFKQIISLESMSNDTSLP 301  
QY 241 DKNSFL 247  
DB 302 DQNSFL 308

## RESULT 9

ID Q23719 PRELIMINARY; PRT; 406 AA.  
AC Q23719;  
DT 01-JAN-1998 (T-EMBLrel. 05, Created)  
DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)  
DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)  
DE MAP3K delta-1 protein kinase (Fragment).  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC Eucosida II; Brassicales; Brassicaceae; Arabidopsids.  
NCBI\_TaxId=3702;  
[1]

## SEQUENCE FROM N.A.

RP STRAIN-COLUMBIA;  
RX MEDLINE=99196996; PubMed=10095117;  
RA Jouanin S., Hamel A., Leprieux A.S., Tregear J.W., Kreis M.,  
RA Henry Y.;  
RT "Characterisation of novel plant genes encoding MEKK/STK1 and RAF-  
RT related protein kinases";  
RL Gene 229:171-181(1999).  
[2]

## SEQUENCE FROM N.A.

RC STRAIN-COLUMBIA;  
RA Jouanin S., Leprieux A.S., Hamel A., Kreis M., Henry Y.;  
RT "Plant MAP kinase signalling pathways in the limekiln";  
RL Adv. Bot. Res. 30:0-0(2000)  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; Y14199; CA474591.1; -  
DR HSSP; P12931; 1PMK.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00221; STYKc; 1.  
DR PROSITE; PSS0107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PSS0011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PSS0108; PROTEIN\_KINASE\_ST; 1.  
KM Kinase; Serine/threonine-protein kinase.  
FT NON TER 1  
SQ SEQUENCE 406 AA; 45965 MW; CE0D94BF2BC12AB CRC64;

Query Match 39.3%; Score 511.5; DB 10; Length 406;  
Best Local Similarity 40.5%; Pred. No. 4.9e-40;  
Matches 104; Conservative 51; Mismatches 85; Indels 17; Gaps 6;

QY 1 YRAKVISQKEVAVVKKLLKIEKAEILSVLSHRNIIQFYGVILEPPNYGIVTEVASLQSL 50  
DB 150 YRAEWM-NGTEVAVVKKLLDQDFSGDALITQKSEILMLRHRNVVLFMGAVRPPNFST 207  
QY 51 VTEYASLSGLYDIYNSNSEMDMHIMTATDVAKGMHYLHMEAPVYVYIHRDLKSRNVV 110  
DB 208 LTFEFLPGSLYRLH-RPNHQLEDEKRRMALDVAKGMHYLHSHPT-VVHRDLKSPNLL 265  
QY 111 IADGVLTICDGFASRFNHT--THMSLVGTFFPMABEVIQSLPVSECTDYSYGVVLA 168  
DB 266 VDGKVVAVKVCDFSLSKMHTIYLSKSTAGTPEWMAPEVLREPAVEKCDVYSFGVILWE 325  
QY 169 MLTREVFPKGLGLQVAVLVEKNERLTITSSCPSPFAELLHQWEADAKRPSFKQIIS 228  
DB 326 LATSRYVFKGLNPMQVAVGFGNRLLEIPDDIDLVQAIIRECHQTEPHLPSFTQLMQ 385  
QY 229 ILESMSNDTSLPDKNS 245  
DB 386 SLKRLQG-LNISNRANT 401

## RESULT 10

ID Q8S9K4 PRELIMINARY; PRT; 880 AA.  
AC Q8S9K4;  
DT 01-JUN-2002 (T-EMBLrel. 21, Created)  
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
DE AT5G11850/FL4P18.20.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC Eucosida II; Brassicales; Brassicaceae; Arabidopsids.  
NCBI\_TaxId=3702;  
[1]  
RP SEQUENCE FROM N.A.  
RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,

RA Bowser L., Carninci P., Chang B., Dale J.M., Goldsmith A.D.,  
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,  
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Saito M.,  
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,  
RA Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,  
RA Ecker J.R.;  
RT "Arabidopsis cDNA clones";  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; A1075641; AAL77650.1; SD9AD50C2B08444A CRC64;  
SQ SEQUENCE 880 AA; 97881 MW; 5D9AD50C2B08444A CRC64;  
Query Match 39.3%; Score 511.5; DB 10; Length 880;  
Best Local Similarity 40.5%; Pred. No. 1.4e-39;  
Matches 104; Conservative 51; Mismatches 85; Indels 17; Gaps 6;  
QY 1 YRAKWSODKEVAVKCL-----KIEKAEILSVLSHRNIOQYGVILLEPPNY 50  
DB 624 YRAEW--NGTEVAVKFKILDODFSGDALTQFKSEIIMLRHPNVVLFMGAVTRPPNFI 681  
D6 51 VTEVASLSGLVDYINSRSEEMDMHMTWATDVAKGMHYLMEAPVYVIRHDLKSRNV 110  
D6 662 LTFELPRGSLVYLH-RENHQIDERRRMALDVAKGMVYLTSHPT-VVHRDLKSPYL 739  
QY 111 IADGVLTICDFGASRPHNHT--TMSLVGTFPMAAPVYQSIPVSECTDYSGVILME 168  
DB 740 VDKMNVKVCDFGLSRMKHHTYLSKSTAGTPENMAPEVLNRPANECVDVSPFVILME 799  
QY 169 MLTRVPEFGLEGLOVAMLVVEKNERLTIPSCPSFAELHOCWEADAKKPSFOIIS 228  
DB 800 IATSRVPMKGLPMQVGVAGVGFQNRRLTLPDIDITVAQIIECQTPHARPSFTQMQ 859  
QY 229 ILESMSNDTSLPDKCNS 245  
DB 860 SLKRLQG-LNISNRANT 875  
RESULT 11  
Q9H2N5 PRELIMINARY; PRT; 1066 AA.  
AC Q9H2N5;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE Mixed lineage kinase MLK1 (Fragment).  
DE Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
NCBI\_TaxID=9606;  
[1]  
SEQUENCE FROM N.A.  
RA McNece J.J., Dower S.K., Guesdon F.;  
RT "cDNA sequence and gene organisation of mixed lineage kinase 1.";  
RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
DR EMBL; AF251442; AAG44591.1; -  
DR HSSP; E29355; ISEM.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001452; SH3.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF00018; SH3; 1.  
DR PRINTS; PR00452; SH3DOMAIN.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR ProDom; PD000066; SH3; 1.  
DR SMART; SM00326; SH3; 1.  
DR SMART; SM00220; S\_TKC; 1.  
DR SMART; SM00219; TYK; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS90108; PROTEIN\_KINASE\_ST; 1.

DR PROSITE; PS50002; SH3; 1.  
KW Kinase; SH3 domain.  
FT NON TER 1  
SQ SEQUENCE 1066 AA; 118463 MW; EDD08BEE7482723 CRC64;  
Query Match 39.1%; Score 508.5; DB 4; Length 1066;  
Best Local Similarity 42.7%; Pred. No. 3.4e-39;  
Matches 108; Conservative 43; Mismatches 77; Indels 25; Gaps 4;  
QY 1 YRAKWSODKEVAVK-----KILKIEKAEILSVLSHRNIOQYGVILLEPPNY 48  
DB 107 YRAFWIGD--EVAVKAARDPDEDISQTIENVRQBAKLFAMKHPNIIALRGVCLKEPVL 164  
QY 49 GIVTEVASLSGLVDYINSRSEEMDMHMTWATDVAKGMHYLMEAPVYVIRHDLKSRN 108  
DB 165 CLMEFARAGGPNLRVLSGR---IPPDILVWVAQIARGMVYLTSHPT-VVHRDLKSSN 221  
QY 109 VVI-----AADGVLTICDFGASRPHNHTTMSLVGTFPMAAPVYQSIPVSECTDY 160  
DB 222 ILIQKVENGDLSNKLKLTIDPGLAREMHTTKMSAAGYAMWADBEVIRASWFSKSDVW 281  
QY 161 SYGVILMEMLTRVPEFGLEGLOVAMLVVEKNERLTIPSCPSFAELHOCWEADAKK 220  
DB 282 SYGVILMEMLTGEVPPRGIDGLAVAGVAMKALPISCTCEPPAKLMEDCWNPDPHSR 341  
QY 221 PSFQIISILESM 233  
DB 342 PSFTNILDQLTTI 354  
RESULT 12  
Q21982 PRELIMINARY; PRT; 740 AA.  
AC Q21982;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE Hypothetical protein R13F6.6.  
GN R13F6.6.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
[1]  
SEQUENCE FROM N.A.  
RA STRAIN=BRISTOL N2;  
RC STRAIN=BRISTOL N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
RT investigating biology. The C. elegans Sequencing Consortium.";  
RT Science 282:2012-2018(1998).  
RL [2]  
SEQUENCE FROM N.A.  
RA STRAIN=BRISTOL N2;  
RT "The sequence of C. elegans cosmid R13F6.";  
RT Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.  
RN [3]  
SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA Waterston R.;  
RT "Direct Submission.";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; U00046; AAC47047.3; -  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00454; SAM; 1.

DR SMART; SM00220; S\_TKC; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE\_ST; 1.  
 KW ATP-binding; Hypothetical protein; Serine/threonine-protein kinase;  
 transferase.  
 KM  
 SEQUENCE 740 AA; 84075 MW; A36B98D1E1D3A0F CRC64;

Query Match 38.1%; Score 495.5; DB 5; Length 740;  
 Best Local Similarity 41.0%; Pred. No. 3.5e-38;  
 Matches 102; Conservative 48; Mismatches 86; Indels 13; Gaps 5;

QY 1 YRAKWSOD---KEVAVKLLKIEKEAEILSVLSHRNIIQFYGVILEPPN-YGIYVEYAS 56  
 DB 67 FSGNMTLPDGSORTIALKVFLEKEAEILSKRHNIIOFYGICATGNDFFIYTEVAE 126  
 QY 57 LGLVYINSNNEEM-----DMDHMTATDVAKGMHYLHMEAPVYIHRDLKSRV 109  
 DB 127 KGLVYPIHSESSQASASSGNSFDVVMASQIASGIQYIHYADVDTIHRDLKSNV 186  
 QY 110 VLAADGLKICDGFASRFNHT-THMSLVGTPEPMAPE-VIOSLPVSECTDYSGVYLM 167  
 DB 187 VLDKMLVCKICDGFSTKDLTHSCTAPSGTAAHMSPEMILQSEGLTATDVMSIGVLM 246  
 QY 168 EMLTEVPKPGLEGLOVAMLVKNERLTPSSCPSPFAELHQCWEADAKRPSFKQII 227  
 DB 247 ELSKEVPEPKDYSEFRITMTQSGITLAIIPSCPAFLKQIMSNCKMTPKRAMRQIQ 306  
 QY 228 SILESMSND 236  
 DB 307 GELNRLAGN 315

## RESULT 13

Q8VDG6 PRELIMINARY; PRT; 1001 AA.

AC Q8VDG6  
 DT 01-MAR-2002 (Tremblrel. 20, Created)  
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Similar to mitogen-activated protein kinase kinase 9.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Strauberg R.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC021891; AAH21891.1;  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR001245; SH3.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR PRODOM; PD000001; Euk\_pkinase; 1.  
 DR PRODOM; PD000066; SH3; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR SMART; SM00220; S\_TKC; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN; 1.  
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN; 1.  
 DR PROSITE; PS00002; SH3; 1.  
 KW Kinase.  
 SQ SEQUENCE 1001 AA; 109983 MW; E10042C686B9953C CRC64;

Query Match 38.0%; Score 494; DB 11; Length 1001;  
 Best Local Similarity 39.4%; Pred. No. 7.3e-38;  
 Matches 108; Conservative 47; Mismatches 85; Indels 34; Gaps 5;

QY 1 YRAKWSODKEVAVKLLK-----IEKEAEILSVLSHRNIIQFYGVILEPPN 48  
 DB 125 YRAWT--OGQVAVAAARDEPDQAAAASVREARLFAALRPNIITQLGVCLRPHL 182  
 QY 49 GIVTEYASLGLVDYI-----NSNREEMDMHMTATDVAKGMHYLHMEAPV 98  
 DB 183 CLVLEFANGALNRLAALAAADPRAPGRARLRIPVOLVMNAVQIRGLYLHEAVVP 242  
 QY 99 VIHRDLKSRNVIT-----ADGVKICDGFASRFNHTTHMSLVGTPEPMAPEYIOS 150  
 DB 243 ILHDLKSNMILILEKIEHDDICNKTILKIDFGIAREHRTTRSAAGTAVMAPEVIRS 302  
 QY 151 LPVSECTDYSGVYVLMEMLTREVPKPGLEGLOVAMLVKNERLTPSSCPSPFAELH 210  
 DB 303 SLFSGKSDIMYGVVLELLELGEVPEYIGDLAAYAVAVANKLTLPISPCPEPFALMK 362  
 QY 211 QCWEADAKRPSFKQIISILESMSND--TSLPDK 242  
 DB 363 ECWEQDPHIRSFALLIQQLTAIEAVLTNMPQE 396

## RESULT 14

Q8WY25 PRELIMINARY; PRT; 859 AA.

AC Q8WY25  
 DT 01-MAR-2002 (Tremblrel. 20, Created)  
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Zipper protein kinase.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Reddy U.R.;  
 RL "Genomic Structure and Promoter Characterization of the Homo sapiens  
 RT MAP3K12 Gene";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF283475; AAL67158.1;  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR PRODOM; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00220; S\_TKC; 1.  
 DR SMART; SM00219; TYRK; 1.  
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN; 1.  
 KW Kinase.  
 SQ SEQUENCE 859 AA; 93219 MW; 1E1BCAD26EDFCF8 CRC64;

Query Match 37.7%; Score 490; DB 4; Length 859;  
 Best Local Similarity 44.3%; Pred. No. 1.4e-37;  
 Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;

QY 10 KEVAVKLLKIEKEAEILSVLSHRNIIQFYGVILEPPNIGIVTEYASLGLVYINSN 67  
 DB 147 EEVAVKRVLDL-KETDIDKRLRLKHPNITTFKGVCTQAPCYILMEFCAGQGLYEVLRAG 205  
 QY 68 RESEMDMHDIMTATDVAKGMHYLHMEAPVYIHRDLKSRNVIAADGVTKICDGSAR- 126  
 DB 206 RPTTPEL-LVMSKGIAGANVYLH--KIHRDLKSPNNLITYDDVVKISDFGTSKE 260  
 QY 127 FNNHTHMSLVGTPEPMAPEVIOSLPVSECTDYSGVYVLMEMLTREVPKPGLEGLOVAM 186  
 DB 261 LSKDSTKMSFAGTAVMAPEVIRNPEVEKVDIWSFGVWELTNGELPYDVDSALIM 320  
 QY 187 LVKNERLTIPSSCPSPFAELHQCWEADAKRPSFKQIISILESMSNDT-SLPDK 242



GenCore version 5.1.4\_p5 4578  
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:12:12 ; Search time 47.5397 Seconds  
(without alignments)  
692.324 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_31\_277

Perfect score: 1300  
Sequence: 1 TRAKISQDKEVAVKKLKI.....SILESMSNTSLPDKNSFL 247

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Actual number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: A\_Geneseq\_101002:\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1300	100.0	455 21 AAB18657	A human regulator
2	1300	100.0	455 21 AAB18657	A human regulator
3	1300	100.0	455 21 AAB18657	A human regulator
4	1300	100.0	455 21 AAB18657	A human regulator
5	1300	100.0	455 21 AAB18657	A human regulator
6	1300	100.0	455 21 AAB18657	A human regulator
7	1300	100.0	455 21 AAB18657	A human regulator
8	1300	100.0	455 21 AAB18657	A human regulator
9	1300	100.0	455 21 AAB18657	A human regulator
10	1300	100.0	455 21 AAB18657	A human regulator

11	490	37.7	859	16	AAB28886	Human leucine zipp
12	490	37.7	859	18	AAB12227	Human leucine-zipp
13	489	37.6	888	23	AAB57049	Mouse ischaemic co
14	487.5	37.5	1021	23	AAB57049	Mouse ischaemic co
15	483	37.2	719	22	AAB85513	Novel human protei
16	483	37.2	1036	23	AAB80923	Human protein kin
17	482.5	37.1	847	23	AAB22763	Novel human protei
18	472.5	36.3	977	22	AAB1694	Human mitogen acti
19	470	36.2	144	22	ABG06092	Drosophila melano
20	469.5	36.1	1490	22	ABG19123	Novel human diagno
21	453	34.8	982	22	AAB50439	Tomato TCTR2. Lyc
22	452	34.8	903	22	AAB50440	Tomato TCTR2. Lyc
23	448.5	34.5	850	23	AAB18529	Rice EDR1. Oriza
24	448	34.5	369	21	AA22172	Melon constitutive
25	448	34.5	374	21	AA22171	Arabidopsis thalia
26	448	34.5	412	21	AA22170	Arabidopsis thalia
27	445.5	34.3	1020	22	AAB58999	Arabidopsis thalia
28	442	34.0	821	16	AAB80574	Drosophila melano
29	442	34.0	821	18	AA17938	Arabidopsis thalia
30	442	34.0	821	22	AAB50438	Constitutive tripl
31	438	33.7	821	15	AA46723	Arabidopsis thalia
32	436.5	33.6	933	22	AAB50437	Arabidopsis thalia
33	435.5	33.5	970	22	AAB50443	Arabidopsis thalia
34	424	32.6	92	22	AA087295	Barley EDR1. Horde
35	424	32.6	92	22	AA087295	Novel central nerv
36	418.5	32.2	589	21	AA45984	Novel signal trans
37	418.5	32.2	732	21	AA45983	Arabidopsis thalia
38	418.5	32.2	760	21	AA45982	Arabidopsis thalia
39	418	32.2	367	21	AA32053	Arabidopsis thalia
40	418	32.2	369	21	AA32052	Arabidopsis thalia
41	418	32.2	407	21	AA32051	Arabidopsis thalia
42	399	30.7	678	22	AAB58061	Arabidopsis thalia
43	395	30.4	567	20	AA28998	Drosophila melano
44	395	30.4	579	18	AA27092	Human TGF-beta act
45	395	30.4	579	18	AA27093	Human transforming

## ALIGNMENTS

RESULT 1		
ID	AAB18657	standard; Protein; 455 AA.
XX		
AC	AAB18657;	
XX		
DT	22-JAN-2001	(first entry)
DE		A human regulator of intracellular phosphorylation.
XX		
KW	Human; intracellular phosphorylation regulator; HRP; stroke; myeloma;	
KW	neurological disorder; Parkinson's disease; demyelinating disease;	
KW	meningitis; developmental disorder; neuromuscular disorder; cancer;	
KW	myasthenia gravis; cell proliferative disorder; actinic keratosis;	
KW	atherosclerosis; atherosclerosis; leukemia; melanoma; bronchitis;	
KW	autoimmune disorder; inflammatory disorder; Addison's disease;	
KW	acquired immunodeficiency disease; allergy; diabetes mellitus;	
KW	rheumatoid arthritis; microbial infection; trauma.	
OS	Homo sapiens.	
XX		
FT		
FT	Key	Location/Qualifiers
FT	Domain	16..257
FT		/note= "eukaryotic protein kinase domain"
FT	Modified-site.	61
FT		/note= "potential phosphorylation site"
FT	Modified-site	89
FT		/note= "potential phosphorylation site"
FT	Modified-site	96
FT		/note= "potential phosphorylation site"
FT	Modified-site	97
FT		/note= "potential glycosylation site"
FT	Binding-site	129..141

FT Modified-site /note= "protein kinase ATP-binding site"  
 159 /note= "potential glycosylation site"  
 FT Modified-site 234 /note= "potential phosphorylation site"  
 FT Modified-site 252 /note= "potential phosphorylation site"  
 FT Modified-site 258 /note= "potential phosphorylation site"  
 FT Modified-site 265 /note= "potential phosphorylation site"  
 FT Modified-site 268 /note= "potential glycosylation site"  
 FT Modified-site 294..322 /note= "potential phosphorylation site"  
 FT Region /note= "leucine zipper"  
 FT Modified-site 302 /note= "potential phosphorylation site"  
 FT Modified-site 302 /note= "potential phosphorylation site"  
 FT Modified-site 342 /note= "potential phosphorylation site"  
 FT Modified-site 343 /note= "potential phosphorylation site"  
 FT Modified-site 346 /note= "potential phosphorylation site"  
 FT Modified-site 364 /note= "potential phosphorylation site"  
 FT Modified-site 409 /note= "potential phosphorylation site"  
 FT Modified-site 410 /note= "potential glycosylation site"  
 FT Modified-site 414 /note= "potential phosphorylation site"  
 FT Modified-site 415 /note= "potential phosphorylation site"  
 FT Modified-site 429 /note= "potential phosphorylation site"  
 FT Modified-site 434 /note= "potential phosphorylation site"  
 FT Modified-site /note= "potential phosphorylation site"  
 PN WO20055332-A2.  
 PD 21-SEP-2000.  
 XX 17-MAR-2000; 2000WO-US07277.  
 XX 18-MAR-1999; 99US-0125593.  
 PR 20-MAY-1999; 99US-0135049.  
 PR 09-JUL-1999; 99US-0143186.  
 (INCY-) INCYTE PHARM INC.  
 XX Bandman O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y;  
 PI Lu DAM, Au-Young J;  
 XX MPI: 2000-602121/57.  
 DR N-PSDB; AA75674.  
 XX Novel human intracellular phosphorylation regulator polypeptides and  
 PT polynucleotides for diagnosis, prevention and treatment of  
 PT neurological, cell proliferative and autoimmune/inflammatory disorders  
 PT  
 PT  
 XX Claim 1; Page 75-76; 96pp; English.  
 PS  
 XX The present sequence represents a human regulator of intracellular  
 CC phosphorylation (HRIP). HRIP is useful for screening agonists and  
 CC antagonists of HRIP polypeptide. HRIP and its agonist or antagonist  
 CC are useful for treating a disease or condition associated with  
 CC decreased or increased expression of functional HRIP. Diseases treated  
 CC or diagnosed include neurological disorders such as stroke, Parkinson's  
 CC disease, demyelinating diseases, bacterial and viral meningitis and  
 CC other developmental disorders of the central nervous system,

CC neuromuscular disorders, myasthenia gravis, cell proliferative disorders  
 CC such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer  
 CC including leukaemia, melanoma, myeloma and cancer of the adrenal gland,  
 CC bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/  
 CC inflammatory disorder such as Addison's disease, acquired  
 CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,  
 CC rheumatoid arthritis, microbial infection and trauma.  
 XX  
 SQ Sequence 455 AA;  
 Query Match 100.0%; Score 1300; DB 21; Length 455;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-127;  
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 YRAKWSIQDEKAVKLLKKEKAEILSVSHNIIQFVGLIEPPNYGLVEYASLSL 60  
 DB 31 YRAKWSIQDEKAVKLLKKEKAEILSVSHNIIQFVGLIEPPNYGLVEYASLSL 90  
 QY 61 YDYINSNRSEMDMIMTATDPAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTIC 120  
 DB 91 YDYINSNRSEMDMIMTATDPAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLTIC 150  
 QY 121 DFGASRFPHNHTTMSLVGTFPPMAPEVYOSLPVSEICDTYSYGVLMEMLTREVPKGL 180  
 DB 151 DFGASRFPHNHTTMSLVGTFPPMAPEVYOSLPVSEICDTYSYGVLMEMLTREVPKGL 210  
 QY 181 GLQVAMLVKRNRLTPSSCPRSFAELLHQCWEADAKKRPFKQIISLESMSNDTSLP 240  
 DB 211 GLQVAMLVKRNRLTPSSCPRSFAELLHQCWEADAKKRPFKQIISLESMSNDTSLP 270  
 QY 241 DKCNSFL 247  
 DB 271 DKCNSFL 277  
 RESULT 2  
 ID AAY83278  
 ID AAY83278 standard; Protein; 455 AA.  
 XX  
 AC AAY83278;  
 XX  
 DT 16-AUG-2000 (first entry)  
 XX  
 DE Human survival regulating kinase (SRK).  
 XX  
 KW Survival regulating kinase; SRK; Raf; MBP; BAD; MAPKK; MAPKK;  
 KW HAX-1; protein kinase; autophosphorylation; cell growth; regulation;  
 KW apoptosis; cell survival; nuclear targeting; tumour; human;  
 KW autoimmune disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200022142-A2.  
 XX  
 PD 20-APR-2000.  
 XX  
 XX 20-SEP-1999; 99WO-US22008.  
 PF  
 XX 13-OCT-1998; 98US-0104088.  
 PR  
 XX (ONIX-) ONYX PHARM INC.  
 PA  
 XX Ruggieri R, Callow M, Diaz P;  
 PI  
 XX MPI: 2000-317994/27.  
 DR N-PSDB; AA293783.  
 XX  
 XX Novel human survival regulating kinase polypeptide for screening agents  
 PT which modulate biological pathways associated with SRK useful in  
 PT treating autoimmune diseases, tumors and apoptosis-related disorders  
 PS  
 XX Claim 4; Figure 2; 62pp; English.



CC Survival regulating kinases (SRK) are a class of proteins involved in  
 CC cell signal transduction pathways such as mitogen-activated protein  
 CC kinase pathways. A protein kinase activity means that the SRK can  
 CC catalyze a reaction in which a phosphate group is transferred from a  
 CC phosphate donor to a phosphate acceptor amino acid residue.  
 CC Substrates for the hydroxyl side chain of a serine or threonine.  
 CC activity is similar to that of a MAPKK such as Raf. has a range of  
 CC other activities including a cell growth-regulatory activity, a cell  
 CC survival promoting activity, a HAX-1 binding activity, an apoptosis  
 CC suppressing activity, a MAPK activation or stimulatory activity, a  
 CC nuclear targeting activity and a SRK-specific immunogenic activity.  
 CC SRK is useful for identifying agents which modulate cellular  
 CC transformations mediated by Ras and SRK and agents that modulate the  
 CC apoptosis suppression activity of SRK. This information may be useful  
 CC in the treatment of autoimmune diseases, tumours and apoptosis  
 CC related disorders.

Query Match 100.0%; Score 1300; DB 21; Length 455;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-127;  
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKMIQDKEVAVKLLKIEKEAEILSVLSHRNIIQFVGVILEPPNYGIVTEVASIGSL 60  
 DB 31 YRAKMIQDKEVAVKLLKIEKEAEILSVLSHRNIIQFVGVILEPPNYGIVTEVASIGSL 90  
 QY 61 YDINSNRSEEMDMHMTATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLIKIC 120  
 DB 91 YDINSNRSEEMDMHMTATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLIKIC 150  
 QY 121 DFGASRFNHTTMSLVGFPPMAPEVIOQLPVSECTDYSGVVLWMLTREVPKGL 180  
 DB 151 DFGASRFNHTTMSLVGFPPMAPEVIOQLPVSECTDYSGVVLWMLTREVPKGL 210  
 QY 181 GLQVAMLVVEKNERLTIPSSCPSPFAELLHQCWEADAKRPSFKQIITILESMSNDTSLP 240  
 DB 211 GLQVAMLVVEKNERLTIPSSCPSPFAELLHQCWEADAKRPSFKQIITILESMSNDTSLP 270  
 QY 241 DKCNSFL 247  
 DB 271 DKCNSFL 277

RESULT 3  
 AAY84321

AAY84321 standard; Protein; 455 AA.

AAY84321;

12-JUL-2000 (first entry)

A human cardiovascular system associated protein kinase-2.

Human; cardiovascular system associated protein kinase-2; CSAPK-2;  
 signalling pathway; cell growth; cell differentiation; gene mapping;  
 tissue typing; forensic identification; cardiovascular disease;  
 congestive heart failure; transgenic animal.

Homo sapiens.

MO200014212-A1.

16-MAR-2000.

09-SEP-1999; 99MO-US20631.

09-SEP-1998; 98US-0099657.  
 29-SEP-1998; 98US-0163115.

(MILL-) MILLENNIUM PHARM INC.

XX

XX

XX

XX

XX

XX

XX

XX

PI Acton S;

DR WPI: 2000-271053/23.  
 DR N-PSDB; AAZ99726, AAZ99727.

PT New nucleic acid encoding cardiovascular system associated protein  
 kinase, used e.g. for diagnosis, treatment and prevention of  
 cardiovascular disease

Claim 2; Fig 2; 163pp; English.

The present sequence represents a human cardiovascular system associated  
 protein kinase-2 (CSAPK-2). CSAPK polypeptides are involved in signalling  
 pathways associated with cell growth and differentiation. The CSAPK  
 polypeptides and polynucleotides are used to screen for agents that  
 specifically modulate CSAPK, which are potential therapeutic agents.  
 They are also used for diagnosis, prognosis or monitoring of  
 CSAPK-related diseases, gene mapping, tissue typing and forensic  
 identification, and for treating or preventing disorders associated  
 with aberrant CSAPK expression or activity, especially cardiovascular  
 diseases such as congestive heart failure. They can also be used in  
 pharmacogenomics. The CSAPK polynucleotide may also be used to generate  
 transgenic animals.

Query Match 100.0%; Score 1300; DB 21; Length 455;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-127;  
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKMIQDKEVAVKLLKIEKEAEILSVLSHRNIIQFVGVILEPPNYGIVTEVASIGSL 60  
 DB 31 YRAKMIQDKEVAVKLLKIEKEAEILSVLSHRNIIQFVGVILEPPNYGIVTEVASIGSL 90  
 QY 61 YDINSNRSEEMDMHMTATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLIKIC 120  
 DB 91 YDINSNRSEEMDMHMTATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVLIKIC 150  
 QY 121 DFGASRFNHTTMSLVGFPPMAPEVIOQLPVSECTDYSGVVLWMLTREVPKGL 180  
 DB 151 DFGASRFNHTTMSLVGFPPMAPEVIOQLPVSECTDYSGVVLWMLTREVPKGL 210  
 QY 181 GLQVAMLVVEKNERLTIPSSCPSPFAELLHQCWEADAKRPSFKQIITILESMSNDTSLP 240  
 DB 211 GLQVAMLVVEKNERLTIPSSCPSPFAELLHQCWEADAKRPSFKQIITILESMSNDTSLP 270  
 QY 241 DKCNSFL 247  
 DB 271 DKCNSFL 277

RESULT 4  
 AAM25322

AAM25322 standard; Protein; 473 AA.

AAM25322;

16-OCT-2001 (first entry)

Human protein sequence SEQ ID NO:837.

Human; cancer; ulcer; HIV infection; human immunodeficiency virus;  
 antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
 antibacterial; endocrine; cardiac; central nervous system; vitinide;  
 anti-HIV; fungicide; antimitogen; cardiovascular; antianemic; anemia;  
 dermatological; antiallergic; antiaesthetic; antidiabetic; cytostatic;  
 neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
 immunostimulant; gene therapy; antitense therapy; vaccine; inflammation;  
 antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;  
 cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
 genetic disease; haematopoietic disorder; platelet disorder; asthma;  
 thrombocytopenia; osteoporosis; severe combined immunodeficiency;

KM		allergic rhinitis; diabetes; multiple sclerosis; depression;
KM		Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
XX		neurological disorder.
OS	Homo sapiens.	
XX	MO200153455-A2.	
PN	26-JUL-2001.	
XX		
PD		
XX	22-DEC-2000; 2000WO-US35017.	
PF		
XX	23-DEC-1999; 99US-0471275.	
PR	21-JAN-2000; 2000US-0488725.	
PR	25-APR-2000; 2000US-0552317.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Tang YT, Liu C, Drmanac RT;	
XX		
XX	WPI: 2001-457603/49.	
XX	N-PSTDB; AAH99263.	
PT		
XX	Isolated human polynucleotides encoding polypeptides, useful for the	
XX	treatment and diagnosis of e.g. cancer, ulcers and HIV infection -	
XX	Claim 20; Page 191; 1217P; English.	
PS		
XX	AAH99166 to AAH99904 encode the human proteins given in AAM25225 to	
CC	AAH995963. The proteins can have activities based on the tissues and	
CC	cells they are expressed in, such as: antiinflammatory; antirheumatic;	
CC	central nervous system; virucide; anti-HIV; fungicide; antimutagen;	
CC	cardiovascular system; antineoplastic; haemostatic; vulnertary;	
CC	antitumor; osteopathic; dermatological; antiallergic; antiaesthetic;	
CC	antidiabetic; cytotoxic; neuroprotective; antidepressant; nootropic;	
CC	antiParkinsonian, and immunostimulant. The proteins and polynucleotides	
CC	encoding them can be used in gene therapy, antisense therapy and vaccine	
CC	production. The proteins and polynucleotides are useful for screening for	
CC	agonists or antagonists of a protein and for the treatment and diagnosis	
CC	of disorders associated with the activity of a protein e.g. inflammation,	
CC	rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,	
CC	neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal	
CC	infections, autoimmunity, genetic diseases, haematopoietic disorders,	
CC	anaemia, platelet disorders, thrombocytopoenia, wounds, burns, ulcers,	
CC	osteoporosis, severe combined immunodeficiency, eczema, allergic	
CC	rinitis, asthma, diabetes, cancer, multiple sclerosis, depression,	
CC	Alzheimer's disease, Parkinson's disease, neurodegenerative and	
CC	neurological disorders.	
XX		
SQ	Sequence 473 AA:	
	Query Match 100.0%; Score 1300; DB 22; Length 473;	
	Best Local Similarity 100.0%; Pred. No. 3,le-127; Indels 0; Gaps 0;	
	Matches 247; Conservative 0; Mismatches 0;	
OY	1 YRAAVTISODKEVAVKLLKTIEKEAEILTSVSHRNIIIOFYGVILLEPPNYGIVTEVASIGSL 60	
Db	49 YRAAVTISODKEVAVKLLKTIEKEAEILTSVSHRNIIIOFYGVILLEPPNYGIVTEVASIGSL 108	
OY	61 YDYINSNRSEEMDMHIWTATVDYAKGMHYLMEAPVKVIHRDIKSRNVITADGYLKIC 120	
Db	109 YDYINSNRSEEMDMHIWTATVDYAKGMHYLMEAPVKVIHRDIKSRNVITADGYLKIC 168	
OY	121 DFGASRRPHNTHHSLSYCTFPFMAPEVYIOSLPVSETCDTISYGVLWMELTREVPFGILE 180	
Db	169 DFGASRRPHNTHHSLSYCTFPFMAPEVYIOSLPVSETCDTISYGVLWMELTREVPFGILE 228	
OY	181 GLQVAVMLVEENKRELTPSSCPRAFALHLQCWEADKKRPSPFOIISLESMSNDTSLP 240	
Db	229 GLQVAVMLVEENKRELTPSSCPRAFALHLQCWEADKKRPSPFOIISLESMSNDTSLP 288	
OY	241 DKCNSFL 247	

[illegible]

QY 1 YRAKISODKEVAVKKLLKIEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEVASLGL 60  
 DB 31 YRAKISODKEVAVKKLLKIEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEVASLGL 90  
 QY 61 YDIYNSNRSEEMDMHIMTATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVKIC 120  
 DB 91 YDIYNSNRSEEMDMHIMTATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVKIC 150  
 QY 121 DFGASRFHNHTHTMSLVGTFFPMMAPEVIOQLPVSETCDTYSYGVVLMEMLTREVPFKGLE 180  
 DB 151 DFGASRFHNHTHTMSLVGTFFPMMAPEVIOQLPVSETCDTYSYGVVLMEMLTREVPFKGLE 210  
 QY 181 GLQVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKKPSFKQIISLESMSNDTSLP 240  
 DB 211 GLQVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKKPSFKQIISLESMSNDTSLP 270  
 QY 241 DKCNSFL 247  
 271 DKCNSFL 277

RESULT 6  
 AAB65673  
 ID AAB65673 standard; Protein; 800 AA.

AC AAB65673;  
 DT 27-MAR-2001 (first entry)

DE Novel protein kinase, SEQ ID NO: 201.

XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;  
 KW immunosuppressive; cardiac; renal; antiinflammatory; antiaesthetic;  
 KW dermatological; antidiabetic; antifertility; gene therapy; vaccine;  
 KW immune disorder; cardiovascular disease; neurodegenerative disease;  
 KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;  
 KW inflammatory pelvic disease; multiple sclerosis; psoriasis.

XX Homo sapiens.  
 PN WO200073469-A2.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US14842.

XX 28-MAY-1999; 99US-0136503.

XX (SUGEN-) SUGEN INC.

XX Plowman GD, Martinez R, Whyte D, Sudersanam S;

XX WPI; 2001-032161/04.

XX DR N-PSDB; AAF44701.

XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
 PT treating immune-related diseases and disorders, cardiovascular disease,  
 PT neurodegenerative diseases and/or cancers -

XX Claim 10; Fig 1; 310pp; English.

XX The present sequence is a novel protein kinase. The novel protein kinases  
 CC and the nucleic acids that encode them may be used in the treatment and  
 CC diagnosis of diseases associated with inappropriate kinase expression  
 CC such as immune-related diseases and disorders, cardiovascular disease,  
 CC neurodegenerative diseases and/or cancers. The nucleic acids and  
 CC complementary sequences may also be used as DNA probes in diagnostic  
 CC assays. The kinase polypeptides may be used as antigens in the production  
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies  
 CC and kinase antagonists may also be used to down regulate kinase  
 CC expression and activity. Diseases related to kinase expression and  
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune

CC disorders, complications of organ transplantation, myocardial infarction,  
 CC immune disorders, cardiovascular diseases, strokes, renal failure,  
 CC oxidative-stress related disorders, chronic inflammatory bowel disease,  
 CC chronic inflammatory pelvic disease, multiple sclerosis, asthma,  
 CC osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and  
 CC reproductive disorders.

XX Sequence 800 AA;

Query Match 100.0%; Score 1300; DB 22; Length 800;

Best Local Similarity 100.0%; Pred. No. 6, 6e-127; Mismatches 0; Gaps 0;

Matches 247; Conservative 0; Indels 0; Gaps 0;

QY 1 YRAKISODKEVAVKKLLKIEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEVASLGL 60  
 DB 31 YRAKISODKEVAVKKLLKIEKEAELISVLSHRNIIQFYGVILEPPNYGIVTEVASLGL 90  
 QY 61 YDIYNSNRSEEMDMHIMTATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVKIC 120  
 DB 91 YDIYNSNRSEEMDMHIMTATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADGVKIC 150  
 QY 121 DFGASRFHNHTHTMSLVGTFFPMMAPEVIOQLPVSETCDTYSYGVVLMEMLTREVPFKGLE 180  
 DB 151 DFGASRFHNHTHTMSLVGTFFPMMAPEVIOQLPVSETCDTYSYGVVLMEMLTREVPFKGLE 210

QY 181 GLQVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKKPSFKQIISLESMSNDTSLP 240  
 DB 211 GLQVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKKPSFKQIISLESMSNDTSLP 270

QY 241 DKCNSFL 247  
 DB 271 DKCNSFL 277

RESULT 7  
 AAG75571  
 ID AAG75571 standard; Protein; 349 AA.

XX AAG75571;

XX 03-SEP-2001 (first entry)

XX Human colon cancer antigen protein SEQ ID NO:6335.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;

XX colorectal carcinoma.

XX Homo sapiens.

XX WO200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26524.

XX 29-SEP-1999; 99US-0157137.

XX 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI; 2001-235357/24.

XX DR N-PSDB; AAH34976.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
 PT useful for preventing, diagnosing and/or treating colorectal cancers -

XX Claim 11; Page 7789-7790; 9803pp; English.  
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where  
 CC the proteins are collectively known as colon cancer antigens. The colon

CC cancer antigens have cytostatic activity and can be used in gene  
 CC therapy and vaccine production. N and P may be used in the prevention,  
 CC diagnosis and treatment of diseases associated with inappropriate P  
 CC expression. For example, N and P may be used to treat disorders  
 CC associated with decreased expression by rectifying mutations or deletions  
 CC in a patient's genome that affect the activity of P by expressing P.  
 CC Inactive proteins or to supplement the patient's own production of P.  
 CC Additionally, N may be used to produce the colon cancer-associated Pg,  
 CC by inserting the nucleic acids into a host cell and culturing the cell  
 CC to express the proteins. N and P can be used in the prevention, diagnosis  
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
 CC and AAB77789 represent sequences used in the exemplification of the  
 CC present invention.  
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
 CC missing at time of publication, meaning no sequences are present for  
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 349 AA;

Query Match 96.4%; Score 1253; DB 22; Length 349;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-122;  
 Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSIDQKVAAYKLIKKEKAEILISVSHRNIIOPGYILEPPNYGIVTEYASIGSL 60  
 DB 87 YRAKWSIDQKVAAYKLIKKEKAEILISVSHRNIIOPGYILEPPNYGIVTEYASIGSL 146  
 QY 61 YDYINSNRSEEMDMHITWATDVAKGMYLHMEAPVKVIRHDKSRNVVIAADGVLIKIC 120  
 DB 147 YDYINSNRSEEMDMHITWATDVAKGMYLHMEAPVKVIRHDKSRNVVIAADGVLIKIC 206  
 QY 121 DFGASRFNNHTTMSLVTFPPMAPEVIOSLPVSECTCTYSYGVLMEMLTREVPFGKLE 180  
 DB 207 DFGASRFNNHTTMSLVTFPPMAPEVIOSLPVSECTCTYSYGVLMEMLTREVPFGKLE 266  
 QY 181 GLQYAMLVENKNERLTIPSSCFRFAELHOCWEADAKKRSPFOIISILSMNSNDTSL 239  
 DB 267 GLQYAMLVENKNERLTIPSSCFRFAELHOCWEADAKKRSPFOIISILSMNSNDTSL 325

RESULT 8  
 AAG03583  
 ID AAG03583 standard; Protein; 141 AA.

XX AAG03583;

DT 06-OCT-2000 (first entry)

Human secreted protein, SEQ ID NO: 7664.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping.

OS Homo sapiens.

XX EPI033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dunas Milne Edwards J, Duclert A, Giordano J;

XX WPI: 2000-500381/45.

XX N-PSDB; AAC03589.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 13; SEQ ID 7664; 71pp + CD-ROM; English.

PS The present sequence is a polypeptide encoded by one of a large number  
 XX of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
 CC different tissues. EST sequences usually correspond mainly to the 3'  
 CC untranslated region (UTR) of the mRNA because they are often obtained  
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 CC those cases where longer cDNA sequences have been obtained, the full 5'  
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 CC ends and can therefore be used to obtain full length cDNAs and genomic  
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 CC chromosome mapping procedures. They are used to obtain upstream  
 CC regulatory sequences and to design expression and secretion vectors.

XX Sequence 141 AA;

Query Match 57.2%; Score 744; DB 21; Length 141;  
 Best Local Similarity 98.6%; Pred. No. 8.8e-70;  
 Matches 139; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 72 MDDMHITWATDVAKGMYLHMEAPVKVIRHDKSRNVVIAADGVLIKICDFGASRFNNHT 131  
 DB 1 MDDMHITWATDVAKGMYLHMEAPVKVIRHDKSRNVVIAADGVLIKICDFGASRFNNHT 60  
 QY 132 THMSLVTFPPMAPEVIOSLPVSECTCTYSYGVLMEMLTREVPFGKLEGLQYAMLVK 191  
 DB 61 THMSLVTFPPMAPEVIOSLPVSECTCTYSYGVLMEMLTREVPFGKLEGLQYAMLVK 120  
 QY 192 NERLTIPSSCFRFAELHOC 212  
 DB 121 NERLTIPSSCFRFAELHOC 141

RESULT 9  
 AAE11775  
 ID AAE11775 standard; Protein; 1046 AA.

XX AAE11775;

DT 18-DEC-2001 (first entry)

XX Human kinase (PKIN)-9 protein.

XX Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;  
 KW cancer; allergy; sarcoma; acquired immune deficiency syndrome;  
 KW AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;  
 KW atherosclerosis; cardiovascular disease; myocardial infarction; anaemia;  
 KW myasthenia gravis; cirrhosis; cataract; growth and development disorder;  
 KW seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;  
 KW lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;  
 KW obesity; restorative therapy; immunomodulatory; vaccine; cardiovascular;  
 KW antimicrobial; cytostatic; antiinflammatory; asthma.

XX Homo sapiens.

XX Location/Qualifiers

XX 55..114 /note="SH3 domain"

XX 134..393 /note="Eukaryotic protein kinase domain"

XX 136..386 /note="Protein kinase domain"

XX 154..207 /note="Receptor tyrosine kinase"

XX 181..228 /note="Receptor tyrosine kinase"

XX 210..223 /note="Tyrosine kinase catalytic site"

XX 232..254 /note="Receptor tyrosine kinase"

Seq	Sequence	1046 AA;
XX	Query Match	38.3%; Score 497.5; DB 22; Length 1046;
XX	Best Local Similarity	42.3%; Pred. No. 1.1e-42;
XX	Matches 107; Conservative	43; Mismatches 78; Indels 25; Gaps 4;
Qy	1 YRAKMIQSDEKVVVK-----KLLKIEKAEILISVLSHNTIQFYGVILEPPNY 48	
Db	149 YRAFMIGD--EVAVKAARHPDEDISQITENVRQEKLPAMLKHPNITIALRGVCLKEPNL 206	
Qy	49 GIYTEVASIGSLYDYNINSRSEMDHIMTWATVAKGMVYLHAEPAKVIHRDLKSRN 108	
Db	207 CLVMEFARAGGPIINRVLSGKR---IPPDILVMNAVQIANGMNYLDEALVPIIHRDLKSN 263	
Qy	109 VVV-----AADGLTKICDFGASRFHHTHTMSLVGFPPMADEVYQSLPVSFCTDY 160	
Db	264 ILLQVENGDSNKLKITDFGLAREWHRTTKMSAQTVAAMPFVIRASMPFSKGDVW 323	
Qy	161 SYGVLMEMLTREVPFKGLEGVQVAMLVVEKNERLTPISSCPSPAEILHQCMEADAKR 220	
Db	324 SYGVLLMELTREVPPRGIDGLAVAVGVAAMKALPIFSTCEPFAKLMEDCKMPDPHSR 383	
Qy	221 PSFKQIISLBSM 233	
Db	384 PSFTNILDQLTTI 396	
RESULT 10		
AAE21717		
ID	AAE21717 standard; Protein, 1097 AA.	
AC	AAE21717;	
XX		
DT	16-JUL-2002 (first entry)	
DE	Human PKIN-12 protein.	
XX		
KW	Human; kinase; enzyme; PKIN-12 protein; immune system disorder; anaemia; acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease; asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease; AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder; Down's syndrome; gene therapy; protein therapy; cytosstatic.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	1..17
FT		/label=Signal_peptide
FT	Protein	18..1097
FT		/note="Mature human PKIN-12 protein"
FT	Domain	55..114
FT		/note="SH3 domain"
FT	Domain	144..403
FT		/note="Eukaryotic protein kinase domain"
FT	Domain	146..396
FT		/note="Protein kinase domain"
FT	Domain	163..396
FT		/note="Protein kinase domain"
FT	Domain	220..233
FT		/note="Tyrosine kinase catalytic domain"
FT	Domain	258..276
FT		/note="Tyrosine kinase catalytic domain"
FT	Domain	311..321
FT		/note="Tyrosine kinase catalytic domain"
FT	Domain	330..352
FT		/note="Tyrosine kinase catalytic domain"
FT	Domain	374..396
FT		/note="Tyrosine kinase catalytic domain"
FT	Domain	438..749
FT		/note="Leucine zipper domain"
FT	Domain	869..893









Db 313 WSYGLWLELLTGVEPVYKIDGLAVAGVAVNKLTLPSTCPPEPAKMKECQGDPHI 372  
 QY 220 RPSFKQIISILESMSND--TSLPDK 242  
 Db 373 RPSFALLIEQLTALIGAVNTMPQE 397

RESULT 15  
 AAB85513

ID AAB85513 standard; protein; 719 AA.

AC AAB85513;

DT 25-SEP-2001 (first entry)

DE Human protein kinase SGK067.

XX Protein kinase; enzyme; cytosolic; noctropic; neuroprotective; human;  
 antiParkinsonian; vincristine; antibacterial; antifungal; antimigraine;  
 analgesic; hypotensive; hypertensive; immunosuppressive; antiallergic;  
 antiparasitic; antirheumatic; antidiabetic; antidiabetic; anorectic;  
 osteopathic; thrombolytic; antiarteriosclerotic; antisthmatic;  
 vasotropic; antidiabetic; gene therapy.

OS Homo sapiens.

PN MO200155356-A2.

PD 02-AUG-2001.

XX 25-JAN-2001; 2001MO-US02337.

XX 25-JAN-2000; 2000US-0178078.

PR 31-JAN-2000; 2000US-0179364.

PR 17-EB-2000; 2000US-0183173.

PR 17-MAR-2000; 2000US-0190162.

PR 29-MAR-2000; 2000US-0193404.

PR 13-NOV-2000; 2000US-0247013.

XX (SUCE-) SUGEN INC.

PI Plowman G, Whyte D, Manning G, Sudareanam S, Martinez R;

XX WPI; 2001-476202/51.

DR N-PSDB; AAH46913.

XX Claim 7; Page 217; 218pp; English.

The invention provides human protein kinases and protein kinase-like enzymes and polynucleotides encoding the polypeptides. The kinase polypeptides and their modulators are useful for treating a disease or disorder such as cancer, immune-related diseases, cardiovascular disease, brain or neuronal-associated disease and metabolic disorders, including cancers of tissues, cancers of hematopoietic origin, diseases of the central nervous system, diseases of the peripheral nervous system, Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis, viral infections, infections caused by prions, bacteria and fungi, ocular diseases, migraines, pain, sexual dysfunction, mood disorders, attention disorders, cognition disorders, hypotension, hypertension, psychotic disorders, neurological disorders, hypotension, metabolic disorders, and organ transplant rejection. They are also useful for treating rhinitis, asthma, chronic inflammatory disease, psoriasis, osteoarthritis, sepsis, chronic inflammatory disease, chronic inflammatory bowel disease, rheumatoid arthritis, metabolic disorders such as diabetes, obesity, cardiovascular diseases such as reperfusion injury, coronary thrombosis, clotting disorders and atherosclerosis, ocular diseases such as glaucoma, retinopathy and macular degeneration, psychiatric and neurological disorders such as anxiety, schizophrenia,

CC dementia, manic depression, etc. The polynucleotides are useful in gene  
 CC therapy techniques to treat the above mentioned disorders. Sequences  
 CC AAB85491-85522 represent the human protein kinases of the invention.  
 XX  
 SQ Sequence 719 AA;

Query Match 37.2%; Score 483; DB 22; Length 719;  
 Best Local Similarity 38.4%; Pred. No. 2.1e-41;  
 Matches 106; Conservative 49; Mismatches 85; Indels 36; Gaps 5;

QY 1 YRAKXISODKEVAVKILK-----IEKAEILSVLSHRNITQFYGVILEPPNY 48  
 Db 139 YRATM--QGQVAVKARQDEPDAAAASVREARLFAMLRHPNIIELRGVCLQDPHL 196  
 QY 49 GIVTEVASLGLVDYI-----NSNRSEMDMDHIMTMDTVAKGMHYLHNEAP 96  
 Db 197 CLVLEFARGALNRLAANAADPPRAGRARRRIPPHVAVMAVOIARGMLYLHNEAP 256  
 QY 97 VKVTHRLKSRNVI-----AADGVLCIDFGASRFPHNHTHNSLVGTFFPMAPEVI 148  
 Db 257 VPIHRLKSSNILLLEKIEHDDICNKTITDPGLAREHRTTKSTAGTVAMAPVI 316  
 QY 149 QSLPVSFETCDTYSYGVVLMELTREVPPKGLBGLQVAMLVENGERLTTPSSCPSPFAL 208  
 Db 317 KSSLFSKSDIMSYGVILWELTGVEPVYKIDGLAVAGVAVNKLTLPSTCPPEPAKL 376  
 QY 209 LHQCWEADAKRPSFKQIISILESMSND--TSLPDK 242  
 Db 377 MKECWODPHIRPSFALLIEQLTALIGAVNTMPQE 412

Search completed: May 1, 2003, 20:35:36  
 Job time : 50.5397 secs



GenCore version 5.1.4\_j5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:38:08 ; Search time 15.8466 Seconds  
(without alignments) 1344.947 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_31\_277

Perfect score: 1300  
Sequence: 1 YRAKWSODKEVAVKKLTKT.....SILESMNDTSLPDKNSFL 247

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328255 seqs, 86286685 residues

a1 number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1300	100.0	455	US-09-757-982-5	Sequence 5, Appl1
2	510.5	39.3	394	US-09-862-027-19	Sequence 19, Appl1
3	495.5	38.1	328	US-09-862-027-18	Sequence 18, Appl1
4	483	37.2	1036	US-10-014-882-2	Sequence 2, Appl1
5	482.5	37.1	847	US-10-143-133-2	Sequence 2, Appl1
6	469.5	36.1	966	US-09-771-161A-197	Sequence 197, App
7	448.5	34.5	850	US-09-904-389-2	Sequence 2, Appl1
8	442	34.0	263	US-09-840-704-5	Sequence 5, Appl1
9	424	32.6	92	US-09-764-868-799	Sequence 799, App
10	395	30.4	579	US-10-158-895-4	Sequence 4, Appl1
11	395	30.4	590	US-10-158-895-15	Sequence 15, Appl1
12	351.5	27.0	251	US-08-987-689A-32	Sequence 32, Appl1
13	343	26.4	505	US-09-977-260-6	Sequence 6, Appl1
14	343	26.4	505	US-09-977-261-6	Sequence 6, Appl1
15	343	26.4	505	US-09-977-266-6	Sequence 6, Appl1
16	343	26.4	505	US-09-982-610-20	Sequence 20, Appl1
17	336	25.8	1036	US-09-771-161A-255	Sequence 255, App
18	336	25.8	1036	US-09-771-161A-256	Sequence 256, App
19	335.5	25.8	537	US-09-977-260-11	Sequence 11, Appl1

20	335.5	25.8	537	9	US-09-977-261-11	Sequence 11, Appl1
21	335.5	25.8	537	10	US-09-977-266-11	Sequence 11, Appl1
22	335.5	25.8	835	10	US-09-947-199-8	Sequence 8, Appl1
23	335	25.8	764	10	US-09-925-302-714	Sequence 714, App
24	333.5	25.7	537	10	US-09-771-161A-212	Sequence 212, App
25	333.5	25.7	537	10	US-09-771-161A-213	Sequence 213, App
26	332	25.5	310	9	US-09-939-833-7	Sequence 7, Appl1
27	332	25.5	310	10	US-09-939-754-7	Sequence 7, Appl1
28	332	25.5	310	10	US-09-939-833-7	Sequence 7, Appl1
29	332	25.5	425	10	US-09-828-313-29	Sequence 29, Appl1
30	332	25.5	822	9	US-09-757-415A-2	Sequence 2, Appl1
31	328.5	25.3	450	9	US-09-977-260-7	Sequence 7, Appl1
32	328.5	25.3	450	9	US-09-977-261-7	Sequence 7, Appl1
33	328.5	25.3	450	10	US-09-977-266-7	Sequence 7, Appl1
34	328.5	25.3	450	10	US-09-947-199-2	Sequence 2, Appl1
35	327.5	25.2	258	10	US-09-840-704-3	Sequence 3, Appl1
36	324.5	25.0	536	9	US-09-977-260-13	Sequence 13, Appl1
37	324.5	25.0	536	9	US-09-929-266-10	Sequence 10, Appl1
38	324.5	25.0	536	9	US-09-977-261-13	Sequence 13, Appl1
39	324.5	25.0	536	10	US-09-977-269-13	Sequence 13, Appl1
40	324	24.9	536	9	US-09-977-260-12	Sequence 12, Appl1
41	324	24.9	536	9	US-09-977-261-12	Sequence 12, Appl1
42	324	24.9	536	10	US-09-977-269-12	Sequence 12, Appl1
43	321	24.7	256	10	US-09-840-704-4	Sequence 4, Appl1
44	320	24.6	543	9	US-09-977-260-14	Sequence 14, Appl1
45	320	24.6	543	9	US-09-977-261-14	Sequence 14, Appl1

## ALIGNMENTS

```
RESULT 1
US-09-757-982-5
; Sequence 5, Application US/09757982
; Patient No. US20020094559A1
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/757,982
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-757-982-5

Query Match      100.0%; Score 1300; DB 10; Length 455;
Best Local Similarity 100.0%; Pred. No. 1e-98;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSODKEVAVKKLTKIEKEAELSVLSHRNIIQFVGLLEPPNYGIYEVASISL 60
DB 31 YRAKWSODKEVAVKKLTKIEKEAELSVLSHRNIIQFVGLLEPPNYGIYEVASISL 90
QY 61 YVYINSNREEDMDHMTWADVAKGMHYLMEAPVYIHRDLKSRNVVIAADGVAKIC 120
DB 91 YVYINSNREEDMDHMTWADVAKGMHYLMEAPVYIHRDLKSRNVVIAADGVAKIC 150
QY 121 DPGASRFNHTHTMSLVGTFFPMAPAEVIOSLPVSECTQTSYGVLAMELTRVFFKLE 180
DB 151 DPGASRFNHTHTMSLVGTFFPMAPAEVIOSLPVSECTQTSYGVLAMELTRVFFKLE 210
QY 181 GLOVAVLVEKXERLTIPSSCRSPFAELLHOCWEADAKKRSFQIISILESMNDTSLP 240
DB 211 GLOVAVLVEKXERLTIPSSCRSPFAELLHOCWEADAKKRSFQIISILESMNDTSLP 270
QY 241 DKCNSFL 247
|||||
```

Db 271 DKNSFL 277

RESULT 2  
US-09-862-027-19  
Sequence 19, Application US/09862027  
Patent No. US20020142428A1  
GENERAL INFORMATION:  
APPLICANT: Hodge, Martin R.  
TITLE OF INVENTION: No. US20020142428A1 Kinases and Uses Thereof  
FILE REFERENCE: 35800/234862  
CURRENT APPLICATION NUMBER: US/09/862,027  
CURRENT FILING DATE: 2001-05-21  
PRIORITY APPLICATION NUMBER: US 09/345,473  
PRIORITY FILING DATE: 1999-06-30  
NUMBER OF SEQ ID NOS: 82  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 19  
LENGTH: 394  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-862-027-19

Query Match 39.3%; Score 510.5; DB 10; Length 394;  
Best Local Similarity 42.7%; Pred. No. 3.2e-34;  
Matches 108; Conservative 44; Mismatches 76; Indels 25; Gaps 4;

QY 1 YRAKMSIQDKEAVK-----KLTKEAEILSVSHRNIIQFYGVILEPPNY 48  
DB 18 YRAFWIGD--EVAVKAARHPEDDISQITENVRQBAKLFAMLKHNIIALGVCCKEENL 75  
QY 49 GIVTEYASISLYDIYINSRSEEMDMHMTATDVAKGMHYLMEAPKYVTHRLKSRN 108  
DB 76 CIVMEFARCGPLNRYLSGKR---IPPDILVNAVOIARGMNYLHDEATVPIIHRDLKSSN 132  
QY 109 VVI-----AADGVLKICDGFASRFNHTTMSLVGTFPPMAPEVILQSLPVSCTDY 160  
DB 133 IILIKVENGDISNKLKLTIDGLAREWHRTTKMSAGTYAMAEVIRASNFSGSDVW 192  
QY 161 SYGVVLMEMLTREVPFKGLEQVAMLVVEKNERLTIPSSCFRSPAEILHQCWEADAKR 220  
DB 193 SYGVILMELLTGEVPRGIDGLRVAGVAMNKLALPIPTCEPEPKLMEDCMNDPSHR 252  
QY 221 PSFKQIISLIESM 233  
DB 253 PSFTNILDQLTTI 265

UNT 3

09-862-027-18  
Sequence 18, Application US/09862027  
Patent No. US20020142428A1  
GENERAL INFORMATION:  
APPLICANT: Hodge, Martin R.  
TITLE OF INVENTION: No. US20020142428A1 Kinases and Uses Thereof  
FILE REFERENCE: 35800/234862  
CURRENT APPLICATION NUMBER: US/09/862,027  
CURRENT FILING DATE: 2001-05-21  
PRIORITY APPLICATION NUMBER: US 09/345,473  
PRIORITY FILING DATE: 1999-06-30  
NUMBER OF SEQ ID NOS: 82  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 18  
LENGTH: 328  
TYPE: PRT  
ORGANISM: C. elegans  
US-09-862-027-18

Query Match 38.1%; Score 495.5; DB 10; Length 328;  
Best Local Similarity 41.0%; Pred. No. 4.4e-33;  
Matches 102; Conservative 48; Mismatches 86; Indels 13; Gaps 5;  
QY 1 YRAKMSIQD--KEVAVKLLIKIEKAEILSVSHRNIIQFYGVILEPPN-YGIVTEYAS 56

Db 67 FSGNWTLPDGSQRTALKKVFLEKEAEILSKIRKHNIIQFYGIKANGNPFITYEAE 126  
QY 57 LGLSLDYINSNRSEEM-----DMDHMTATDVAKGMHYLMEAPKYVTHRLKSRNV 109  
DB 127 KQSLYDFIHSSESSQSPASSGNSFDVVVKMAISOIASGIQYLAHYAVVTIITHRLKSRNV 186  
QY 110 VVADGVKICDGFASRFNHT--TMSLVGTFPPMAPE-VIOSLPVSETCDYSGVVLW 167  
DB 187 VLDKNIIVCKICDGFYSKDIITHSCTAPSWGTYAMNSPEKILQSLTATDVWISGVVLW 246  
QY 168 EMLTREVPFKGLEQVAMLVVEKNERLTIPSSCFRSPAEILHQCWEADAKRPSFKQII 227  
DB 247 EILSKVEPKYKDISEPRIFMTIQSGITLAIPPSCPAFLKQLSNCKMTPKDRAWNRQIQ 306  
QY 228 SLIESMSND 236  
DB 307 GELNRLAGN 315

RESULT 4  
US-10-014-882-2  
Sequence 2, Application US/10014882  
Patent No. US20020107384A1  
GENERAL INFORMATION:  
APPLICANT: Kieke, James  
APPLICANT: Donoho, Gregory  
TITLE OF INVENTION: No. US20020107384A1 Human Kinase and Polynucleotides Encoding  
FILE REFERENCE: LEX-0279-USA  
CURRENT APPLICATION NUMBER: US/10/014,882  
CURRENT FILING DATE: 2001-12-11  
PRIORITY APPLICATION NUMBER: US 60/254,744  
PRIORITY FILING DATE: 2000-12-11  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 1036  
TYPE: PRT  
ORGANISM: homo sapiens  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (1)...(1036)  
OTHER INFORMATION: Xaa = Any Amino Acid  
US-10-014-882-2

Query Match 37.2%; Score 483; DB 12; Length 1036;  
Best Local Similarity 38.4%; Pred. No. 1.7e-31;  
Matches 106; Conservative 49; Mismatches 85; Indels 36; Gaps 5;

QY 1 YRAKMSIQDKEAVKLLK-----IEKAEILSVSHRNIIQFYGVILEPPNY 48  
DB 139 YRAFW--QCGEVAVKARQDEPDAAAASVREARLFLMRLRPNTIIERGVCLQDPHL 196  
QY 49 GIVTEYASISLYIYI-----NSNRSEEMDMHMTATDVAKGMHYLMEAP 96  
DB 197 CIVMEFARCGALNRLAANAAPDPRAPGPRARRIPHYLVNAVOIARGMNYLHDEAF 256  
QY 97 VKVTHRLKSRNVYI-----AADGVLKICDGFASRFNHTTMSLVGTFPPMAPEVI 148  
DB 257 VPIIHRDLKSSNIIILEKIEDHDIQNTKLTIDGLAREWHRTTKMSAGTYAMAEVIR 316  
QY 149 QSLPVSCTDYSGVVLWMLTREVPFKGLEQVAMLVVEKNERLTIPSSCFRSPAEIL 208  
DB 317 KSLSPKSDIWSXGVVLMLTGEVPRGIDGLAVVAVVAVNKLTIPIPTCEPEPKL 376  
QY 209 LHQCWEADAKRPSFKQIISLIESMSND--TSLPDK 242  
DB 377 MKECWQDPHIRPSFALITLBOITALEGAVMTMEMFOE 412

RESULT 5  
US-10-143-133-2

```

; Sequence 2, Application US/10143133
; Publication No. US20020197658A1
; GENERAL INFORMATION:
; APPLICANT: Yoganathan, Thillaiathan
; APPLICANT: Delaney, Allen
; TITLE OF INVENTION: Cancer Associated Protein Kinase and Its Use
; FILE REFERENCE: KIN-023
; CURRENT APPLICATION NUMBER: US/10/143,133
; CURRENT FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: 60/290,555
; PRIOR FILING DATE: 2001-05-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-143-133-2

```

```

Query Match 37.1%; Score 482.5; DB 9; Length 847;
Best Local Similarity 41.1%; Pred. No. 1.5e-31;
Matches 104; Conservative 44; Mismatches 80; Indels 25; Gaps 6;

```

```

QY 1 YRAKISQ-----DKAVAKKLIKIEKAEIISVLSHRNIIQFYGVILEPPNYG 49
DB 132 YKSGRGEIVAVKARQDEEDISV-TESVROEARLPMLHPNIALKAVCLEEPNLC 190
QY 50 IYEVASIGSLVDYINSNRSEEMDMHIM-TWATDVAKGMHYLHMEAPVKYIHRDLKSRN 108
DB 191 LMEVYAAGGPIRSLAGRRVP---PHVLVMAVQIARGMYLHCEALVPYIHRDLKSN 246
QY 109 VV---IAADV---LKICDFGASRFNNHTHMSLVGTFPMMADEVIOQLPVSECTDTY 160
DB 247 ILLLOPISSDDHEHKTLLITDFGLAREWHKTTQMSAGTAMMAPEVIKASTFSGSDHW 306
QY 161 SYGVLMEMLTREVPFKGLEGLQVAMLVKNERLTIPSSCRSPFAELHQCWEADAKR 220
DB 307 SFGVLMELTLGVEVYRIGIDCLAVAYGVAVNKLTPISCTCEPPAQLMADCMQDPHRR 366
QY 221 PSFKQIISLESMS 233
DB 367 PDFASILLQOLEAL 379

```

## RESULT 6

```

US-09-771-161A-197
; Sequence 197, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 197
; LENGTH: 966
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-197

```

```

Query Match 36.1%; Score 469.5; DB 10; Length 966;
Best Local Similarity 42.2%; Pred. No. 2e-30;
Matches 97; Conservative 47; Mismatches 77; Indels 9; Gaps 5;
QY 10 KEVAVKKLIEKEAEI--LSVLSHRNIIQFYGVILEPPNYGVITEVASYISGLVDYINSN 67

```

```

DB 190 EEAIAIK-VREQNETDIKHLAKLKHPIILAFKGCCTQAPCCIIIMEYCAIGQILEVIRAG 248
QY 68 RSEEMDMHMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVVIAADGVKICDFGASR- 126
DB 249 R-KITPRLVDMSTGIASGNNVYLH--KIHRDLKSRNVVLTHTDAKISDFGSK 303
QY 127 FNNHTHMSLVGTFPMMAPEVIOQLPVSECTDTYSGVLMEMLTREVPFKGLEGLQVAM 186
DB 304 LSDSTYMSFAGVAMMAPEVIRNEPVSEKDIWSFGVLMELTLGIEIPYKDVSSAIIW 363
QY 187 LVKNERLTIPSSCRSPFAELHQCWEADAKRPSFKQIISLESMSND 236
DB 364 GVGNSLSLHPVPSTCTCPDGFILMKQTWGSKPRRNPSPRTLMHLDIASAD 413

```

## RESULT 7

```

US-09-904-389-2
; Sequence 2, Application US/09904389
; Patent No. US20020129404A1
; GENERAL INFORMATION:
; APPLICANT: Clendenen, Stephanie K.
; APPLICANT: Schuster, Debra K.
; TITLE OF INVENTION: CTRL HOMOLOGUE FROM MELON
; FILE REFERENCE: 4257-0029.30
; CURRENT APPLICATION NUMBER: US/09/904,389
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/218,307
; PRIOR FILING DATE: 2000-07-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 850
; TYPE: PRT
; ORGANISM: Cucumis melo
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (154) ..(154)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-904-389-2

```

```

Query Match 34.5%; Score 448.5; DB 10; Length 850;
Best Local Similarity 40.2%; Pred. No. 9e-29;
Matches 99; Conservative 38; Mismatches 94; Indels 15; Gaps 4;

```

```

QY 1 YRAKWSIDKEVAVKLIKIE-----KEAELISVLSHRNIIQFYGVILEPPNYGI 50
DB 592 YKGEWHGSD--VAVKLTLEQDPHPRVNEFLREVAIMKSLRHPNIVLPFGAVTKPPNLSI 649
QY 51 VTEVASYISGLVDYINSNRSEEMDMHMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNV 110
DB 650 VTEVYSRGSLLYLLHKSQVKDIDETRIMADPAKGMHYLHRRDP-PIVHRDLSPNLL 708
QY 111 IAAADVLIKICDFGASRFNNHT--THMSLVGTFPMMAPEVIOQLPVSECTDTYSGVLM 168
DB 709 VPKKYVAVKCDGFLSKARTFLSSKSAAGTPEMMAPEVLRDEPNESKSDVYSFGVILME 768
QY 169 MLTREVPFKGLEGLQVAMLVKNERLTIPSSCRSPFAELHQCWEADAKRPSFKQIIS 228
DB 769 IATLQPFMCNLPQVAAVAGFGKRLDIPDVNPKLASLIACWADEPWKRKSPSSIME 828
QY 229 ILESMS 234
DB 829 TLKPMT 834

```

## RESULT 8

```

US-09-840-704-5
; Sequence 5, Application US/09840704
; Patent No. US20020122801A1
; GENERAL INFORMATION:
; APPLICANT: Debnar, Shoukat
; APPLICANT: Hamigan, Greg

```

Query Match	32.6%	Score 424	DB 9	Length 92
Best Local Similarity	98.8%	Pctd. No. 7.3e-28		
Matches 80	Conservative 1	Mismatches 0	Indels 0	Gaps 0
QY 23	EALTLSTLSRHNTIQFVGLLEPPNYGLIVEEYASIGSLYDINSNRSEEDMDHIMTWAT	82		

```

RESULT 11
US-10-158-895-15
; Sequence 15, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158, 895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529, 279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796

```

PRIOR FILING DATE: 1998-10-22  
 PRIOR APPLICATION NUMBER: JP 9/290188  
 PRIOR FILING DATE: 1997-10-22  
 NUMBER OF SEQ ID NOS: 48  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 15  
 LENGTH: 590  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-158-895-15

Query Match  
 Best Local Similarity 37.1%; Score 395; DB 9; Length 590;  
 Matches 91; Conservative 46; Mismatches 80; Indels 28; Gaps 11;

QY 2 RAKWISQDEKAVAVKLLKIEKAE-----ILSLSHRNIIOPGYILEPPNYGIYT 52  
 DB 52 KAM--RAKDAVK--QIESESEKAFIVELRQLSRVNHPVVKLYGACLP--VCLVM 104  
 53 EYASIGSLVDYINSRSEEM--DMDHMTATDVAKGMHYLMEAPVKVIHRDLKSRNV 109  
 105 EYASIGSLVDYINSRSEEM--DMDHMTATDVAKGMHYLMEAPVKVIHRDLKSRNV 162  
 QY 110 VIAADG-VLKICDPG-ASRPHNHTTMSLVGTPEPMAPREVIGSLPSETCDTYSYGVULM 167  
 DB 163 LVAAGTVALKICDPGACIOHTMTNK--GSAAMAPREVIGSYSEKCVFSGIILM 220  
 QY 168 EMLTREVPRKLEG--LOVAVLVEKNERLTIPSSCPSPFAELHQCWEADAKRPSFKQ 225  
 DB 221 EYITRRKPEDEIGCAPFRIM-AVHNGTRPPLIKLPRKISLMRCMSKQPSQPSNEE 279  
 QY 226 ITSL 230  
 DB 280 IVKIM 284

RESULT 12  
 US-09-987-689A-32  
 Sequence 32, Application US/08987689A  
 Patent No. US20020048782A1  
 GENERAL INFORMATION:  
 APPLICANT: Sima Lev  
 APPLICANT: Joseph Schlessinger  
 TITLE OF INVENTION: PYK2 RELATED PRODUCTS AND METHODS  
 NUMBER OF SEQUENCES: 32  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Lyon & Lyon  
 STREET: 633 West Fifth Street  
 CITY: Los Angeles  
 STATE: California  
 COUNTRY: USA  
 ZIP: 90071  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 MB storage  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: IBM P.C. DOS 5.0  
 SOFTWARE: FastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/987,689A  
 FILING DATE: December 9, 1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/032,824  
 FILING DATE: December 11, 1996  
 APPLICATION NUMBER: 08/460,626  
 FILING DATE: June 2, 1995  
 APPLICATION NUMBER: 08/357,642  
 FILING DATE: December 15, 1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Warburg, Richard J.  
 REGISTRATION NUMBER: 32,327  
 REFERENCE/DOCKET NUMBER: 230/110  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600  
 TELEFAX: (213) 955-0440  
 TELE: 67-3510  
 INFORMATION FOR SEQ ID NO: 32:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 251 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-987-689A-32

Query Match  
 Best Local Similarity 32.9%; Score 351.5; DB 8; Length 251;  
 Matches 78; Conservative 50; Mismatches 94; Indels 15; Gaps 7;

QY 1 YRAKWSQDEKAVAVKLL-----LKIE--KEAEILSVLSHRNIIOPGYILEPPNYGIYT 53  
 DB 15 YEGVWKYSILTVAVKTLKEDTMEVEBELKEAAMKEIKHNVLQGLGCTREPPFIITTE 74  
 QY 54 YASIGSLVDYINSRSEEMDMHMTATDVAKGMHYLMEAPVKVIHRDLKSRNVIAA 113  
 DB 75 EMTYGNLIDYLRCCNRQEVAVVLLVWATQISSAMEYLEKK--NFIHRDLAARNCLVGE 131  
 QY 114 DGVLIKIDFGASRPHNHTTMSLVGT-FP--NMAPREVIGSLPSETCDTYSYGVULMEL 170  
 DB 132 NMLVAVDFGLSLMGDTITTAHAKFPIKWTAPSLAVNKFISKSDVAFGLLMEIA 191  
 QY 171 TREY-PRKLEGLOVAVLVEKNERLTIPSSCPSPFAELHQCWEADAKRPSFKQI 226  
 DB 192 TYGWSYPGIDRSQV-YELLEKDYRMKRECGCEKXYELMRACWQNPSPDRSFAEI 247

RESULT 13  
 US-09-977-260-6  
 Sequence 6, Application US/09977260  
 Publication No. US20020192790A1  
 GENERAL INFORMATION:  
 APPLICANT: ULLRICH, AXEL  
 APPLICANT: GISHIZKY, MIKHAEL  
 APPLICANT: SURES, IRMINGARD  
 TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
 FILE REFERENCE: 038602/1260  
 CURRENT APPLICATION NUMBER: US/09/977,260  
 CURRENT FILING DATE: 2001-10-16  
 PRIOR APPLICATION NUMBER: 08/232,545  
 PRIOR FILING DATE: 1994-04-22  
 NUMBER OF SEQ ID NOS: 24  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 6  
 LENGTH: 505  
 TYPE: PRT  
 ORGANISM: Unknown Organism  
 FEATURE:  
 OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte  
 OTHER INFORMATION: Kinase 3  
 US-09-977-260-6

Query Match  
 Best Local Similarity 33.7%; Score 343; DB 9; Length 505;  
 Matches 85; Conservative 50; Mismatches 95; Indels 22; Gaps 9;

QY 12 VAVKLL-----KIEKAEILSVLSHRNIIOPGYV-ILEPPNYGIYTEVASIGSLDY 63  
 DB 259 VAVKTLKPGSDPNDPLREAGIMKNLHPKIQLYAVCTLEDPIY-IITELMRGSLQY 317  
 QY 64 INSRSEEMDMHMTATDVAKGMHYLMEAPVKVIHRDLKSRNVIAADGVLIKIDFG 123  
 DB 318 LQNDTGSKIHLLTQGVDAQAQVAGAAVLESR--NYIHRDLAARNVAVGEHNIVKADFG 374  
 QY 124 ASRPHN-----HTTMSLVGTPEPMAPREVIGSLPSETCDTYSYGVULMELT-REVPF 176  
 DB 375 LARVFKVNDIEDYESRHEITLPPVKTAPPAIRSNKSIKSDVWSFGILLVETIYTGMPY 434

QY 177 KGLEGLQVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKRPSFKQIISILES-MSN 235  
DB 435 SGMGTGAQVIOQLMAQ-NYRLPQPSNCPQOQFYNIWLCWNAEPKRPFTFETLRWKLDEYFET 493  
QY 236 DTSLPDKNSFL 247  
DB 494 DSSYSD-ANNFI 504

## RESULT 14

US-09-977-261-6  
Sequence 6, Application US/09977261  
Publication No. US20030054527A1  
GENERAL INFORMATION:  
APPLICANT: ULTRICH, AXEL  
APPLICANT: GISHIZKY, MIKHAIL  
APPLICANT: SURES, IRMINGARD  
TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
FILE REFERENCE: 038602/1259  
CURRENT APPLICATION NUMBER: US/09/977,261  
CURRENT FILING DATE: 2001-10-16  
PRIOR APPLICATION NUMBER: 08/232,545  
PRIOR FILING DATE: 1994-04-22  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 505  
TYPE: PRT  
ORGANISM: Unknown Organism  
FEATURE:  
OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte  
OTHER INFORMATION: Kinase 3  
US-09-977-261-6

Query Match 26.4%; Score 343; DB 9; Length 505;  
Best Local Similarity 33.7%; Pred. No. 2.1e-20;  
Matches 85; Conservative 50; Mismatches 95; Indels 22; Gaps 9;

QY 12 VAVKTL-----KIEKEAETLSVLSHRNIIQFYGV-ILPEPNYGIIVTEASIGSLYDY 63  
DB 259 VAVKTLKPSMPDNDFLRBAQIMKILRHPKLIQVAVCTLEDPIY-ITTELMRHGSLQEY 317  
QY 64 INSNSSEEMDMHIMTWATDVAKGWHYLMFAPVKVIRHDLKSRNVVIAADGVLKICDFG 123  
DB 318 LQNDTSGKTHLQOVMAQVASGMAVLESR---NYIHRDLAARVVLGEHNITYVADFG 374  
QY 124 ASRFHN-----HTHMSLVGTFPMWAPVIOQLPVSETCDTYSYGVVLMEMLT-REVPF 176  
DB 375 LARVFNVEDIEDYESHEIKLPVKWTAPPAIRSNKFSIKSDVWSFGILLYEITIGKMPY 434  
QY 177 KGLEGLQVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKRPSFKQIISILES-MSN 235  
DB 435 SGMGTGAQVIOQLMAQ-NYRLPQPSNCPQOQFYNIWLCWNAEPKRPFTFETLRWKLDEYFET 493  
QY 236 DTSLPDKNSFL 247  
DB 494 DSSYSD-ANNFI 504

## RESULT 15

US-09-977-269-6  
Sequence 6, Application US/09977269  
Patent No. US20020082037A1  
GENERAL INFORMATION:  
APPLICANT: ULTRICH, AXEL  
APPLICANT: GISHIZKY, MIKHAIL  
APPLICANT: SURES, IRMINGARD  
TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES  
FILE REFERENCE: 038602/1260  
CURRENT APPLICATION NUMBER: US/09/977,269  
CURRENT FILING DATE: 2001-10-16  
PRIOR APPLICATION NUMBER: 08/232,545

PRIOR FILING DATE: 1994-04-22  
NUMBER OF SEQ ID NOS: 24  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 505  
TYPE: PRT  
ORGANISM: Unknown Organism  
FEATURE:  
OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte  
OTHER INFORMATION: Kinase 3  
US-09-977-269-6

Query Match 26.4%; Score 343; DB 10; Length 505;  
Best Local Similarity 33.7%; Pred. No. 2.1e-20;  
Matches 85; Conservative 50; Mismatches 95; Indels 22; Gaps 9;

QY 12 VAVKTL-----KIEKEAETLSVLSHRNIIQFYGV-ILPEPNYGIIVTEASIGSLYDY 63  
DB 259 VAVKTLKPSMPDNDFLRBAQIMKILRHPKLIQVAVCTLEDPIY-ITTELMRHGSLQEY 317  
QY 64 INSNSSEEMDMHIMTWATDVAKGWHYLMFAPVKVIRHDLKSRNVVIAADGVLKICDFG 123  
DB 318 LQNDTSGKTHLQOVMAQVASGMAVLESR---NYIHRDLAARVVLGEHNITYVADFG 374  
QY 124 ASRFHN-----HTHMSLVGTFPMWAPVIOQLPVSETCDTYSYGVVLMEMLT-REVPF 176  
DB 375 LARVFNVEDIEDYESHEIKLPVKWTAPPAIRSNKFSIKSDVWSFGILLYEITIGKMPY 434  
QY 177 KGLEGLQVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKRPSFKQIISILES-MSN 235  
DB 435 SGMGTGAQVIOQLMAQ-NYRLPQPSNCPQOQFYNIWLCWNAEPKRPFTFETLRWKLDEYFET 493  
QY 236 DTSLPDKNSFL 247  
DB 494 DSSYSD-ANNFI 504

Search completed: May 1, 2003, 20:52:14  
Job time: 17.8466 secs



GenCore version 5.1.4.g5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using bw model

Run on: May 1, 2003, 20:33:12 ; Search time 16.5356 seconds  
(without alignments)  
439.505 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_31\_277

Sequence: 1 YRAKWSODKEVAVKLLKLT.....SLESMSNDTSLPDKNSFL 247

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1300	100.0	455	3 US-09-221-235-5	Sequence 5, App1
2	1300	100.0	455	3 US-09-221-928-5	Sequence 5, App1
3	1300	100.0	455	4 US-09-221-527-5	Sequence 5, App1
4	1300	100.0	455	4 US-09-221-236-5	Sequence 5, App1
5	1300	100.0	455	4 US-09-221-416-5	Sequence 5, App1
6	1300	100.0	455	4 US-09-221-245-5	Sequence 5, App1
7	1300	100.0	455	4 US-09-221-115-5	Sequence 5, App1
8	1300	100.0	455	4 US-09-221-528-5	Sequence 5, App1
9	1300	100.0	455	4 US-09-593-553-5	Sequence 5, App1
10	1300	100.0	455	4 US-09-221-237-5	Sequence 5, App1
11	490	37.7	859	1 US-08-205-018-2	Sequence 2, App1
12	490	37.7	859	1 US-08-395-580-2	Sequence 2, App1
13	490	37.7	859	1 US-08-395-580-2	Sequence 2, App1
14	490	37.7	859	1 US-08-395-580-2	Sequence 2, App1
15	490	37.7	859	1 US-08-395-580-2	Sequence 2, App1
16	490	37.7	859	1 US-08-395-580-2	Sequence 2, App1
17	490	37.7	859	1 US-08-395-580-2	Sequence 2, App1
18	490	37.7	859	1 US-08-395-580-2	Sequence 2, App1
19	490	37.7	859	1 US-08-395-580-2	Sequence 2, App1
20	490	37.7	859	1 US-08-395-580-2	Sequence 2, App1
21	490	37.7	859	1 US-08-395-580-2	Sequence 2, App1
22	490	37.7	859	1 US-08-395-580-2	Sequence 2, App1
23	490	37.7	859	1 US-08-395-580-2	Sequence 2, App1
24	490	37.7	859	1 US-08-395-580-2	Sequence 2, App1
25	490	37.7	859	1 US-08-395-580-2	Sequence 2, App1
26	490	37.7	859	1 US-08-395-580-2	Sequence 2, App1
27	490	37.7	859	1 US-08-395-580-2	Sequence 2, App1

28	343	26.4	505	1 US-08-222-616-20	Sequence 20, App1
29	343	26.4	505	4 US-08-446-648-20	Sequence 20, App1
30	343	26.4	505	4 PCT-US95-04228-20	Sequence 20, App1
31	341	26.2	506	4 US-08-426-509A-6	Sequence 6, App1
32	341	26.2	511	5 PCT-US95-05008-6	Sequence 6, App1
33	339	26.1	820	1 US-08-166-717D-6	Sequence 6, App1
34	335.5	25.8	259	2 US-07-857-224B-52	Sequence 52, App1
35	335.5	25.8	537	4 US-08-426-509A-11	Sequence 11, App1
36	335.5	25.8	537	5 PCT-US95-05008-11	Sequence 11, App1
37	335	25.8	313	1 US-08-278-089A-17	Sequence 17, App1
38	335	25.8	313	2 US-08-838-957A-16	Sequence 16, App1
39	335	25.8	729	1 US-07-640-029-3	Sequence 3, App1
40	335	25.8	731	1 US-07-921-807B-5	Sequence 5, App1
41	335	25.8	731	1 US-08-441-944A-5	Sequence 5, App1
42	335	25.8	731	4 US-08-439-992A-3	Sequence 3, App1
43	335	25.8	733	1 US-07-640-029-4	Sequence 4, App1
44	335	25.8	733	1 US-07-921-807B-6	Sequence 6, App1
45	335	25.8	733	1 US-08-441-944A-6	Sequence 6, App1

## ALIGNMENTS

```
RESULT 1
US-09-221-235-5
; Sequence 5, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,235
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-235-5

Query Match      100.0%; Score 1300; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. 1.6e-128;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSODKEVAVKLLKLTKEKAEILSVLSHRNIIQFGVILEPPNYGIVTEYASLGL 60
    |||||
DB 31 YRAKWSODKEVAVKLLKLTKEKAEILSVLSHRNIIQFGVILEPPNYGIVTEYASLGL 90
    |||||
QY 61 YDYNRSSEMDMDHMTWATDVAKGMHYLMHEAVKVIHSDLKSRNVVIAADGVLKIC 120
    |||||
DB 91 YDYNRSSEMDMDHMTWATDVAKGMHYLMHEAVKVIHSDLKSRNVVIAADGVLKIC 150
    |||||
QY 121 DFGASFNHHTHTMSLVGTFPMWAPREVISLPVSECTDYSGVVLWMLTREVPFKGLE 180
    |||||
DB 151 DFGASFNHHTHTMSLVGTFPMWAPREVISLPVSECTDYSGVVLWMLTREVPFKGLE 210
    |||||
QY 181 GLOVAALVYKNERLTPSSCRPSFAELHQCWEADAKRPSFKQIISLESMSNDTSLP 240
    |||||
DB 211 GLOVAALVYKNERLTPSSCRPSFAELHQCWEADAKRPSFKQIISLESMSNDTSLP 270
    |||||
QY 241 DKCNSTL 247
    |||||
DB 271 DKCNSTL 277
    |||||

RESULT 2
US-09-221-928-5
; Sequence 5, Application US/09221928
; Patent No. 6121030
; GENERAL INFORMATION:
```

```

APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/221,928
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE:
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PRT
ORGANISM: Homo sapiens
US-09-221-928-5

Query Match          100.0%; Score 1300; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 1.6e-128;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSIQDEKAVAVKKLIKIEKAEILSVLSHRNIIQFYGVILPEPNYGIETEVASLGS 60
DB 31 YRAKWSIQDEKAVAVKKLIKIEKAEILSVLSHRNIIQFYGVILPEPNYGIETEVASLGS 90

QY 61 YDYINSNRSEEMDMHMTATVAKGMHYLHMEAPVKVIRHDKSRNVVIAADGVLTIC 120
DB 91 YDYINSNRSEEMDMHMTATVAKGMHYLHMEAPVKVIRHDKSRNVVIAADGVLTIC 150

QY 121 DFGASRFNHTTHMSLVGTFFPMAPEVIOQLPVSETCDITYSYGVLMEMLTREVPFKGLE 180
DB 151 DFGASRFNHTTHMSLVGTFFPMAPEVIOQLPVSETCDITYSYGVLMEMLTREVPFKGLE 210

QY 181 GLOVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSPKOIISLESMSNDTSLP 240
DB 211 GLOVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSPKOIISLESMSNDTSLP 270

QY 241 DKCNSFL 247
DB 271 DKCNSFL 277

RESULT 3
US-09-221-527-5
; Sequence 5, Application US/09221527
; Patent No. 6146832
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,527
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-527-5

Query Match          100.0%; Score 1300; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 1.6e-128;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSIQDEKAVAVKKLIKIEKAEILSVLSHRNIIQFYGVILPEPNYGIETEVASLGS 60
DB 31 YRAKWSIQDEKAVAVKKLIKIEKAEILSVLSHRNIIQFYGVILPEPNYGIETEVASLGS 90

QY 61 YDYINSNRSEEMDMHMTATVAKGMHYLHMEAPVKVIRHDKSRNVVIAADGVLTIC 120
DB 91 YDYINSNRSEEMDMHMTATVAKGMHYLHMEAPVKVIRHDKSRNVVIAADGVLTIC 150

QY 121 DFGASRFNHTTHMSLVGTFFPMAPEVIOQLPVSETCDITYSYGVLMEMLTREVPFKGLE 180
DB 151 DFGASRFNHTTHMSLVGTFFPMAPEVIOQLPVSETCDITYSYGVLMEMLTREVPFKGLE 210

QY 181 GLOVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSPKOIISLESMSNDTSLP 240
DB 211 GLOVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSPKOIISLESMSNDTSLP 270

QY 241 DKCNSFL 247
DB 271 DKCNSFL 277

RESULT 4
US-09-221-236-5
; Sequence 5, Application US/09221236
; Patent No. 6146841
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,236
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-236-5

Query Match          100.0%; Score 1300; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 1.6e-128;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSIQDEKAVAVKKLIKIEKAEILSVLSHRNIIQFYGVILPEPNYGIETEVASLGS 60
DB 31 YRAKWSIQDEKAVAVKKLIKIEKAEILSVLSHRNIIQFYGVILPEPNYGIETEVASLGS 90

QY 61 YDYINSNRSEEMDMHMTATVAKGMHYLHMEAPVKVIRHDKSRNVVIAADGVLTIC 120
DB 91 YDYINSNRSEEMDMHMTATVAKGMHYLHMEAPVKVIRHDKSRNVVIAADGVLTIC 150

QY 121 DFGASRFNHTTHMSLVGTFFPMAPEVIOQLPVSETCDITYSYGVLMEMLTREVPFKGLE 180
DB 151 DFGASRFNHTTHMSLVGTFFPMAPEVIOQLPVSETCDITYSYGVLMEMLTREVPFKGLE 210

QY 181 GLOVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSPKOIISLESMSNDTSLP 240
DB 211 GLOVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSPKOIISLESMSNDTSLP 270

QY 241 DKCNSFL 247
DB 271 DKCNSFL 277

RESULT 5
US-09-221-416-5
; Sequence 5, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
```

```

DB 151 DFGASRFNHTTHMSLVGTFFPMAPEVIOQLPVSETCDITYSYGVLMEMLTREVPFKGLE 210
QY 181 GLOVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSPKOIISLESMSNDTSLP 240
DB 211 GLOVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSPKOIISLESMSNDTSLP 270

QY 241 DKCNSFL 247
DB 271 DKCNSFL 277

RESULT 4
US-09-221-236-5
; Sequence 5, Application US/09221236
; Patent No. 6146841
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,236
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-236-5

Query Match          100.0%; Score 1300; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 1.6e-128;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKWSIQDEKAVAVKKLIKIEKAEILSVLSHRNIIQFYGVILPEPNYGIETEVASLGS 60
DB 31 YRAKWSIQDEKAVAVKKLIKIEKAEILSVLSHRNIIQFYGVILPEPNYGIETEVASLGS 90

QY 61 YDYINSNRSEEMDMHMTATVAKGMHYLHMEAPVKVIRHDKSRNVVIAADGVLTIC 120
DB 91 YDYINSNRSEEMDMHMTATVAKGMHYLHMEAPVKVIRHDKSRNVVIAADGVLTIC 150

QY 121 DFGASRFNHTTHMSLVGTFFPMAPEVIOQLPVSETCDITYSYGVLMEMLTREVPFKGLE 180
DB 151 DFGASRFNHTTHMSLVGTFFPMAPEVIOQLPVSETCDITYSYGVLMEMLTREVPFKGLE 210

QY 181 GLOVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSPKOIISLESMSNDTSLP 240
DB 211 GLOVAMLVVEKNERLTIPSSCPSPFAELHQCWEADAKKRPSPKOIISLESMSNDTSLP 270

QY 241 DKCNSFL 247
DB 271 DKCNSFL 277

RESULT 5
US-09-221-416-5
; Sequence 5, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
```

TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-221-416-5

Query Match 100.0%; Score 1300; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1.6e-128;  
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKVISODKEVAVKLLKIEKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 60  
DB 31 YRAKVISODKEVAVKLLKIEKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 90  
QY 61 YDIYNSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVVIAADGVLTIC 120  
DB 91 YDIYNSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVVIAADGVLTIC 150  
QY 121 DFGASRFPHNHTTMSLVGTFFPMMAPEVIQSLPVSETCDITYSYGVLMWMLTREVPFKGLE 180  
DB 151 DFGASRFPHNHTTMSLVGTFFPMMAPEVIQSLPVSETCDITYSYGVLMWMLTREVPFKGLE 210  
DB 181 GLOVAMLVVEKNERLITPSSCFRFAELHQCWEADAKKPSFKQIISLESMSNDTSLP 240  
QY 211 GLOVAMLVVEKNERLITPSSCFRFAELHQCWEADAKKPSFKQIISLESMSNDTSLP 270  
QY 241 DKCNSFL 247  
DB 271 DKCNSFL 277

RESULT 6  
US-09-221-245-5  
Sequence 5, Application US/09221245  
Patent No. 6180358  
GENERAL INFORMATION:  
APPLICANT: Acton, Susan  
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
FILE REFERENCE: NMI-050  
CURRENT APPLICATION NUMBER: US/09/221,245  
CURRENT FILING DATE: 1998-12-28  
EARLIER APPLICATION NUMBER: US 09/163,115  
EARLIER FILING DATE: 1998-09-29  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 455  
TYPE: PRT  
ORGANISM: Homo sapiens  
-09-221-245-5

Query Match 100.0%; Score 1300; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1.6e-128;  
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKVISODKEVAVKLLKIEKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 60  
DB 31 YRAKVISODKEVAVKLLKIEKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 90  
QY 61 YDIYNSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVVIAADGVLTIC 120  
DB 91 YDIYNSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVVIAADGVLTIC 150  
QY 121 DFGASRFPHNHTTMSLVGTFFPMMAPEVIQSLPVSETCDITYSYGVLMWMLTREVPFKGLE 180  
DB 151 DFGASRFPHNHTTMSLVGTFFPMMAPEVIQSLPVSETCDITYSYGVLMWMLTREVPFKGLE 210  
QY 181 GLOVAMLVVEKNERLITPSSCFRFAELHQCWEADAKKPSFKQIISLESMSNDTSLP 240  
DB 211 GLOVAMLVVEKNERLITPSSCFRFAELHQCWEADAKKPSFKQIISLESMSNDTSLP 270  
QY 241 DKCNSFL 247  
DB 271 DKCNSFL 277

RESULT 7  
US-09-163-115-5  
Sequence 5, Application US/09163115A  
Patent No. 6183962  
GENERAL INFORMATION:  
APPLICANT: Acton, Susan  
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
FILE REFERENCE: NMI-050  
CURRENT APPLICATION NUMBER: US/09/163,115A  
CURRENT FILING DATE: 1998-09-29  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 455  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-163-115-5

Query Match 100.0%; Score 1300; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1.6e-128;  
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKVISODKEVAVKLLKIEKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 60  
DB 31 YRAKVISODKEVAVKLLKIEKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 90  
QY 61 YDIYNSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVVIAADGVLTIC 120  
DB 91 YDIYNSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVVIAADGVLTIC 150  
QY 121 DFGASRFPHNHTTMSLVGTFFPMMAPEVIQSLPVSETCDITYSYGVLMWMLTREVPFKGLE 180  
DB 151 DFGASRFPHNHTTMSLVGTFFPMMAPEVIQSLPVSETCDITYSYGVLMWMLTREVPFKGLE 210  
QY 181 GLOVAMLVVEKNERLITPSSCFRFAELHQCWEADAKKPSFKQIISLESMSNDTSLP 240  
DB 211 GLOVAMLVVEKNERLITPSSCFRFAELHQCWEADAKKPSFKQIISLESMSNDTSLP 270  
QY 241 DKCNSFL 247  
DB 271 DKCNSFL 277

RESULT 8  
US-09-221-528-5  
Sequence 5, Application US/09221528  
Patent No. 6190874  
GENERAL INFORMATION:  
APPLICANT: Acton, Susan  
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
FILE REFERENCE: NMI-050  
CURRENT APPLICATION NUMBER: US/09/221,528  
CURRENT FILING DATE: 1998-12-28  
EARLIER APPLICATION NUMBER: 09/163,115  
EARLIER FILING DATE: 1998-09-29  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 455  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-221-528-5

Query Match 100.0%; Score 1300; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1.6e-128;  
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YRAKVISODKEVAVKLLKIEKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 60  
DB 31 YRAKVISODKEVAVKLLKIEKEAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGL 90  
QY 61 YDIYNSNRSEEMDMHIMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVVIAADGVLTIC 120

Db 91 YDIYNSRSEEMDMHITMTATDVAKGMHYLHMEAPVKVHRDLKSRNVVIADGVLTIC 150  
|||  
Qy 121 DFGASRFHNTHTMSLVGTTPMMAPEVIOQLPVSETCDTYSYGVVLMEMLTREVPFKGLE 180  
|||  
Db 151 DFGASRFHNTHTMSLVGTTPMMAPEVIOQLPVSETCDTYSYGVVLMEMLTREVPFKGLE 210  
|||  
Qy 181 GLOVAMLVVEKNERLTIPSSCFPSFAELLHQCWEADAKKRPSPFOIISILESMSNDTSLP 240  
|||  
Db 211 GLOVAMLVVEKNERLTIPSSCFPSFAELLHQCWEADAKKRPSPFOIISILESMSNDTSLP 270  
|||  
Qy 241 DKCNSFL 247  
|||  
Db 271 DKCNSFL 277

RESULT 9  
US-09-593-553-5  
; Sequence 5, Application US/09593553  
; Patent No. 6200770

## GENERAL INFORMATION:

APPLICANT: Accon, Susan  
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
FILE REFERENCE: NMI-050  
CURRENT APPLICATION NUMBER: US/09/593,553  
CURRENT FILING DATE: 2000-06-14  
PRIOR APPLICATION NUMBER: 09/163,115  
PRIOR FILING DATE: 1998-09-28  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 5  
LENGTH: 455  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-593-553-5

Query Match 100.0%; Score 1300; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1,66-128;  
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YRAKWSIQDEKVAVKLLKTEKEAELISVSHRNIIQFYGVILEPPNYGIVTEYASLSL 60  
|||  
Db 31 YRAKWSIQDEKVAVKLLKTEKEAELISVSHRNIIQFYGVILEPPNYGIVTEYASLSL 90  
|||  
Qy 61 YDIYNSRSEEMDMHITMTATDVAKGMHYLHMEAPVKVHRDLKSRNVVIADGVLTIC 120  
|||  
Db 91 YDIYNSRSEEMDMHITMTATDVAKGMHYLHMEAPVKVHRDLKSRNVVIADGVLTIC 150  
|||  
Qy 121 DFGASRFHNTHTMSLVGTTPMMAPEVIOQLPVSETCDTYSYGVVLMEMLTREVPFKGLE 180  
|||  
Db 151 DFGASRFHNTHTMSLVGTTPMMAPEVIOQLPVSETCDTYSYGVVLMEMLTREVPFKGLE 210  
|||  
Qy 181 GLOVAMLVVEKNERLTIPSSCFPSFAELLHQCWEADAKKRPSPFOIISILESMSNDTSLP 240  
|||  
Db 211 GLOVAMLVVEKNERLTIPSSCFPSFAELLHQCWEADAKKRPSPFOIISILESMSNDTSLP 270  
|||  
Qy 241 DKCNSFL 247  
|||  
Db 271 DKCNSFL 277

RESULT 10  
US-09-221-237-5  
; Sequence 5, Application US/09221237  
; Patent No. 6214597  
; GENERAL INFORMATION:  
APPLICANT: Accon, Susan  
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
FILE REFERENCE: NMI-050  
CURRENT APPLICATION NUMBER: US/09/221,237  
CURRENT FILING DATE: 1998-12-28  
PRIOR APPLICATION NUMBER: 09/163,115  
EARLIER FILING DATE: 1998-09-29

NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-237-5

Query Match 100.0%; Score 1300; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1,66-128;  
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YRAKWSIQDEKVAVKLLKTEKEAELISVSHRNIIQFYGVILEPPNYGIVTEYASLSL 60  
|||  
Db 31 YRAKWSIQDEKVAVKLLKTEKEAELISVSHRNIIQFYGVILEPPNYGIVTEYASLSL 90  
|||  
Qy 61 YDIYNSRSEEMDMHITMTATDVAKGMHYLHMEAPVKVHRDLKSRNVVIADGVLTIC 120  
|||  
Db 91 YDIYNSRSEEMDMHITMTATDVAKGMHYLHMEAPVKVHRDLKSRNVVIADGVLTIC 150  
|||  
Qy 121 DFGASRFHNTHTMSLVGTTPMMAPEVIOQLPVSETCDTYSYGVVLMEMLTREVPFKGLE 180  
|||  
Db 151 DFGASRFHNTHTMSLVGTTPMMAPEVIOQLPVSETCDTYSYGVVLMEMLTREVPFKGLE 210  
|||  
Qy 181 GLOVAMLVVEKNERLTIPSSCFPSFAELLHQCWEADAKKRPSPFOIISILESMSNDTSLP 240  
|||  
Db 211 GLOVAMLVVEKNERLTIPSSCFPSFAELLHQCWEADAKKRPSPFOIISILESMSNDTSLP 270  
|||  
Qy 241 DKCNSFL 247  
|||  
Db 271 DKCNSFL 277

RESULT 11  
US-08-205-018-2  
; Sequence 2, Application US/08205018  
; Patent No. 5554523

## GENERAL INFORMATION:

APPLICANT: Reddy, Usharani R.  
TITLE OF INVENTION: Pleasure, David  
TITLE OF INVENTION: No. 5554523el Protein Kinase, Nucleic Acid  
TITLE OF INVENTION: Sequences Encoding the Same and Methods Related  
TITLE OF INVENTION: Thereeto  
NUMBER OF SEQUENCES: 2

## CORRESPONDENCE ADDRESSES:

ADDRESSER: Woodcock Washburn Kurtz Mackiewicz and  
ADDRESSER: No. 5554523ris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/205,018

## FILING DATE:

CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:

NAME: Gaumond, Rebecca R.  
REGISTRATION NUMBER: 35,152  
REFERENCE/DOCKET NUMBER: CH-0488  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:

LENGTH: 668 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

\*  
MOLECULE TYPE: protein  
US-08-205-018-2

Query Match 37.7%; Score 490; DB 1; Length 668;  
Best Local Similarity 44.3%; Pred. No. 4.9e-43;  
Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;

QY 10 KEVAVKLLIKIEKAEI--LSVLSHRNIIQFYGVILEPPNYGVITEVYASLSGLYDYNIN 67  
DB 147 EEVAVKVRDL-KETDIKRLKLPNIITFGVCTQAPCCIMFECAQGLTEVLRAG 205  
QY 68 RSEEMDMHMTATDVAKGMHYLMEAPVYIHRDLKSRNVVIAADGVLCIDFGASR- 126  
DB 206 RPYTPSL--LVDSMGIGAMNYLHLH--KIHRDLKSPNMLITYDDVVKISDFGTSKE 260  
QY 127 FHNHTHMSLVGTFFPMAPEVYIQLPVSECTDYSGVYLMEMLTRVFPFGLEGLOVAV 186  
DB 261 LSDKSTKMSFAGTVAMAPVYIRNEPVSEKDIWSFGVLMELLTGELPYKDVSSAIIW 320

187 LVEKNERLTPSSCPSPFALLHQCWEADAKKPSFKQIISLESNSNDT-SLPDK 242  
321 GVGNSLHLPVSSCPDGFKILLRQCWNKSRNRPSPFQILLHLDIASAVLSTPGE 377

RESULT 12  
US-08-395-580-2

Sequence 2, Application US/08395580  
Patent No. 5676945

GENERAL INFORMATION:  
APPLICANT: Usharani R. Reddy, David Pleasure and the Children's  
APPLICANT: Hospital of Philadelphia  
TITLE OF INVENTION: No. 5676945el Protein Kinase, Nucleic Acid  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Woodcock Washburn Kurtz Mackiewicz and No. 5676945r1e  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch disk, 720 KB  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/395, 580  
FILING DATE: herewith  
CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/205, 018  
FILING DATE: 01-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Rebecca L. Ralph (formerly Gaumond)  
REGISTRATION NUMBER: 35,152  
REFERENCE/DOCKET NUMBER: CH-0488  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 859 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-395-580-2

Query Match 37.7%; Score 490; DB 1; Length 859;  
Best Local Similarity 44.3%; Pred. No. 7e-43;  
Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;

QY 10 KEVAVKLLIKIEKAEI--LSVLSHRNIIQFYGVILEPPNYGVITEVYASLSGLYDYNIN 67

DB 147 EEVAVKVRDL-KETDIKRLKLPNIITFGVCTQAPCCIMFECAQGLTEVLRAG 205  
QY 68 RSEEMDMHMTATDVAKGMHYLMEAPVYIHRDLKSRNVVIAADGVLCIDFGASR- 126  
DB 206 RPYTPSL--LVDSMGIGAMNYLHLH--KIHRDLKSPNMLITYDDVVKISDFGTSKE 260  
QY 127 FHNHTHMSLVGTFFPMAPEVYIQLPVSECTDYSGVYLMEMLTRVFPFGLEGLOVAV 186  
DB 261 LSDKSTKMSFAGTVAMAPVYIRNEPVSEKDIWSFGVLMELLTGELPYKDVSSAIIW 320

187 LVEKNERLTPSSCPSPFALLHQCWEADAKKPSFKQIISLESNSNDT-SLPDK 242  
321 GVGNSLHLPVSSCPDGFKILLRQCWNKSRNRPSPFQILLHLDIASAVLSTPGE 377

RESULT 13  
PCT-US95-02792-2

Sequence 2, Application PC/TUS9502792

GENERAL INFORMATION:  
APPLICANT: Usharani R. Reddy, David Pleasure and the Children's  
APPLICANT: Hospital of Philadelphia  
TITLE OF INVENTION: Novel Protein Kinase, Nucleic Acid  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Woodcock Washburn Kurtz Mackiewicz and Norris  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch disk, 720 KB  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/02792  
FILING DATE: herewith  
CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/205, 018  
FILING DATE: 01-MAR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Rebecca L. Ralph (formerly Gaumond)  
REGISTRATION NUMBER: 35,152  
REFERENCE/DOCKET NUMBER: CH-0488  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 859 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-02792-2

Query Match 37.7%; Score 490; DB 5; Length 859;  
Best Local Similarity 44.3%; Pred. No. 7e-43;  
Matches 105; Conservative 43; Mismatches 79; Indels 10; Gaps 6;

QY 10 KEVAVKLLIKIEKAEI--LSVLSHRNIIQFYGVILEPPNYGVITEVYASLSGLYDYNIN 67  
DB 147 EEVAVKVRDL-KETDIKRLKLPNIITFGVCTQAPCCIMFECAQGLTEVLRAG 205  
QY 68 RSEEMDMHMTATDVAKGMHYLMEAPVYIHRDLKSRNVVIAADGVLCIDFGASR- 126  
DB 206 RPYTPSL--LVDSMGIGAMNYLHLH--KIHRDLKSPNMLITYDDVVKISDFGTSKE 260  
QY 127 FHNHTHMSLVGTFFPMAPEVYIQLPVSECTDYSGVYLMEMLTRVFPFGLEGLOVAV 186

Db 261 LSDSTKMSFAGTAVAMAPVINEPVEKVDIWSFGVILMELLTGEIPYKOVDSALIM 320  
 QY 187 LVYEKNERLTIPSSCPSPFAELLHQCEWADAKRPSFKQIISLESMSNDT-SLPRK 242  
 Db 321 GVSNSLHLFVPSCCPDGFKILLRQCWNSKPRNRPSFRQILHLHDIASADVJSTPGE 377

## RESULT 14

US-09-035-706-5  
 ; Sequence 5, Application US/09035706  
 ; Patent No. 6001622

GENERAL INFORMATION:  
 ; APPLICANT: Dedhat, Shoukat

APPLICANT: Hannigan, Greg  
 ; TITLE OF INVENTION: Integrin-Linked Kinase and  
 ; NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Bozicevic & Reed, LLP  
 ; STREET: 285 Hamilton Avenue, Suite 200  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA

ZIP: 94301  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/035,706  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:  
 ; FILING DATE:  
 ; REGISTRATION NUMBER: 36,677  
 ; REFERENCE/DOCKET NUMBER: KIN-2C1P1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 650-327-3400  
 ; TELEFAX: 650 327-3231

TELEX:  
 ; INFORMATION FOR SEQ ID NO: 5:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 263 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein

US-09-035-706-5  
 ; Query Match 34.0%; Score 442; DB 3; Length 263;  
 ; Best Local Similarity 40.2%; Pred. No. 1,4e-38;  
 ; Matches 99; Conservative 41; Mismatches 90; Indels 16; Gaps 5;

QY 1 YRAKKISODKEVAVAKLKIE-----KEAEILSVLSHRNIIQFYGVILEPPNYGI 50  
 Db 21 HRAEWHGSD--VAVKILMEQDFHAEVRVNEFLREVAIMKRLRHPNIVLFMGAVVQPPMLSI 78  
 QY 51 VTEVASLGLSDYIN-SNRSEEMDMHMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNV 109  
 Db 79 VTEYISRGSLVRLHKSARGQLDERRLSWAYDVAKGMNVLHNRNP-PIVHRDLKSPNL 137  
 QY 110 VIAADGVLIKDFGASRPHNHT--THMSLVGTFFPMMAPEVIOSLPVSETCDYISYGVVLM 167  
 Db 138 LVDKKYTVAVCDPFGSLRKASTFLSSKSAAGTPEMMADEVLRDEPSNKSIVYSFGVILW 197  
 QY 168 EMULTEVPEPKGLEQVAMLVVEKNERLTIPSSCPSPFAELLHQCEWADAKRPSFKQII 227  
 Db 198 ELATLQQPWGNINPQVAAVGFCKRLEIPENLNPQVAAIIEGCTWNEPWKRBSFATIM 257

QY 228 SILESM 233  
 Db 258 DILRPL 263

RESULT 15  
 US-08-955-841-5  
 ; Sequence 5, Application US/08955841  
 ; Patent No. 6013782

GENERAL INFORMATION:  
 ; APPLICANT: Dedhat, Shoukat

APPLICANT: Hannigan, Greg  
 ; TITLE OF INVENTION: Integrin-Linked Kinase and  
 ; NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Bozicevic & Reed, LLP  
 ; STREET: 285 Hamilton Avenue, Suite 200  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA

ZIP: 94301  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/955,841  
 ; FILING DATE:  
 ; CLASSIFICATION: 800  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:  
 ; FILING DATE:  
 ; REGISTRATION NUMBER: 36,677  
 ; REFERENCE/DOCKET NUMBER: KIN-2C1P1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 650-327-3400  
 ; TELEFAX: 650 327-3231

TELEX:  
 ; INFORMATION FOR SEQ ID NO: 5:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 263 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein

US-08-955-841-5  
 ; Query Match 34.0%; Score 442; DB 3; Length 263;  
 ; Best Local Similarity 40.2%; Pred. No. 1,4e-38;  
 ; Matches 99; Conservative 41; Mismatches 90; Indels 16; Gaps 5;

QY 1 YRAKKISODKEVAVAKLKIE-----KEAEILSVLSHRNIIQFYGVILEPPNYGI 50  
 Db 21 HRAEWHGSD--VAVKILMEQDFHAEVRVNEFLREVAIMKRLRHPNIVLFMGAVVQPPMLSI 78  
 QY 51 VTEVASLGLSDYIN-SNRSEEMDMHMTWATDVAKGMHYLHMEAPVKYIHRDLKSRNV 109  
 Db 79 VTEYISRGSLVRLHKSARGQLDERRLSWAYDVAKGMNVLHNRNP-PIVHRDLKSPNL 137  
 QY 110 VIAADGVLIKDFGASRPHNHT--THMSLVGTFFPMMAPEVIOSLPVSETCDYISYGVVLM 167  
 Db 138 LVDKKYTVAVCDPFGSLRKASTFLSSKSAAGTPEMMADEVLRDEPSNKSIVYSFGVILW 197  
 QY 168 EMULTEVPEPKGLEQVAMLVVEKNERLTIPSSCPSPFAELLHQCEWADAKRPSFKQII 227  
 Db 198 ELATLQQPWGNINPQVAAVGFCKRLEIPENLNPQVAAIIEGCTWNEPWKRBSFATIM 257  
 QY 228 SILESM 233

Sun May 4 10:01:21 2003

Db 258 DLRL 263

Search completed: May 1, 2003, 20:39:41  
Job time : 28.5356 secs

us-09-757-982-5\_copy\_31\_277.ra1





GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using 6w model

Run on: May 1, 2003, 20:31:42 ; Search time 0.899582 Seconds

(without alignment)  
1602.986 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_407\_421

Perfect score: 74

Sequence: 1 KONSSTTSKRGRKK 15

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

283224

a1 number of hits satisfying chosen parameters:

283224

Minimum DB seq length: 0

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: Dirl.\*  
2: Dirl.\*  
3: Dirl.\*  
4: Dirl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	62.2	216	2 D75567	GTP cyclohydrolase
2	46	62.2	530	2 F86467	hypothetical prote
3	45	60.8	857	2 A42861	protein kinase TTK
4	44	59.5	425	2 S58672	extra sex combs pr
5	43	58.1	683	2 T34103	hypothetical prote
6	42	56.8	292	2 G88448	protein C45G9.4 [1
7	42	56.8	294	2 E88448	protein C45G9.9 [1
8	42	56.8	312	2 S56289	regulatory protein
9	42	56.8	374	2 T19910	hypothetical prote
10	42	56.8	752	2 S64750	probable ATP-depen
11	42	56.8	1946	2 JC6032	lactocepin (EC 3.4
12	41	55.4	453	2 G96695	hypothetical prote
13	41	55.4	895	2 T34308	hypothetical prote
14	40	54.1	208	2 S43434	histone H1, testic
15	40	54.1	229	2 C45345	vif protein - capr
16	40	54.1	298	2 S63238	hypothetical prote
17	40	54.1	560	2 C71621	hypothetical domain
18	40	54.1	1233	2 S56271	exonuclease
19	39	52.7	81	2 A41949	DNA topoisomerase
20	39	52.7	107	2 H84658	hypothetical prote
21	39	52.7	319	2 AC3455	UDP-galactose-4-epi
22	39	52.7	401	2 T42655	hypothetical prote
23	39	52.7	441	2 T02191	nucleosidase prote
24	39	52.7	546	2 T25257	hypothetical prote
25	39	52.7	667	2 A40713	cylicin I - bovine
26	39	52.7	673	2 BVECB	excinuclease ABC C
27	39	52.7	673	2 A99736	excision nuclease
28	39	52.7	673	2 B85586	DNA repair, excisi
29	39	52.7	769	2 F89870	serine proteinase

30	39	52.7	905	2 A54654	centromere protein
31	39	52.7	952	2 S64473	translation initia
32	39	52.7	1481	1 QZD03	pyrimidine synthe
33	39	52.7	1485	1 ISZPT2	DNA topoisomerase
34	39	52.7	1791	2 T02345	hypothetical prote
35	39	52.7	1898	2 T42440	phospholipase C ho
36	39	52.7	1922	2 T21581	hypothetical prote
37	39	52.7	3869	2 A48205	All-1 protein +GTR
38	38	52.0	1079	2 T38913	translation initia
39	38	51.4	80	2 F84090	hypothetical prote
40	38	51.4	142	2 T51651	midline homolog -
41	38	51.4	142	2 JC4272	pleiotrophic facto
42	38	51.4	142	2 JC4168	midline precursor
43	38	51.4	142	2 JC4273	pleiotrophic facto
44	38	51.4	258	2 S69056	histone H1 - yeast
45	38	51.4	293	2 B89802	conserved hypoche

#### ALIGNMENTS

##### RESULT 1

D75567  
GTP cyclohydrolase I - Deinococcus radiodurans (strain R1)

C/Species: Deinococcus radiodurans

C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000

C/Accession: D75567

R/White, O.; Bisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A/Reference number: A75250; PMID:20036986; PMID:10567266

A/Accession: D75567

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-216 <WHI>

A/Cross-references: GB:AE001867; GB:AE000513; NID:G6457693; PIDN:AAF09628.1; PID:G64576

A/Experimental source: strain R1

C/Genetics:

A/Map position: 1

C/Superfamily: GTP cyclohydrolase I

Query Match 62.2% Score 46; DB 2; Length 216;

Best Local Similarity 76.9%; Pred No. 3.5;

Matches: 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KONSSTTSKRGR 13

DB 181 KONSSTTSAMRG 193

##### RESULT 2

F86467  
hypothetical protein F7P12.5 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001

C/Accession: F86467

R/Rheologis, A.; Becker, J.R.; Palm, C.J.; Federpriel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Corn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luhrs, J.S.; Malt, R.; Matzfall,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A66141; MUID:21016719; PMID:11130712

A/Accession: F86467

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-530 <STO>  
 A/Cross-references: GB:AE005172; NID:G10092384; PIDN:AA612791.1; GSPDB:GN00141  
 C/Genetics:  
 A/Map position: 1

Query Match 62.2%; Score 46; DB 2; Length 530;  
 Best Local Similarity 53.3%; Pred. No. 7.4;  
 Matches 9; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

OY 1 KONSSTSKRRGKK 15  
 Db 460 EDDTKTKTKRGRK 474

## RESULT 3

A42861  
 protein kinase TTK (EC 2.7.1.-) - human  
 N/Alternate names: phosphotyrosine pickled threonine kinase (PYT)  
 C/Species: Homo sapiens (man)  
 C/Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 03-Nov-2000

Accession: A42861; S27971; 138144  
 J. Biol. Chem. 267, 16000-16006, 1992  
 A/Title: Expression of TTK, a novel human protein kinase, is associated with cell prolif

A/Reference number: A42861; MUID:92348472; PMID:1639825  
 A/Accession: A42861

A/Molecule type: mRNA  
 A/Residues: 1-857 <MIL1>  
 A/Cross-references: EMBL:M86699; NID:G340010

A/Note: sequence extracted from NCBI backbone (NCBIN:109875, NCBI:P.109876)  
 A/Note: it is uncertain whether Met-1 or Met-17 is the initiator  
 R/Mills, G.B.; Schmandt, R.; McGill, M.; Amendola, A.; Hill, M.; Jacobs, K.; May, C.; Ro

submitted to the EMBL Data Library, February 1992  
 A/Description: Expression of TTK, a novel human protein kinase.  
 A/Reference number: S27971  
 A/Accession: S27971

A/Molecule type: mRNA  
 A/Residues: 17-857 <MTL2>  
 A/Cross-references: EMBL:M86699; NID:G340010; PIDN:AAA61239.1; PID:G340011

R/Andberg, R.A.; Fischer, W.H.; Hunter, T.  
 Oncogene 8, 351-359, 1993  
 A/Title: Characterization of a human protein threonine kinase isolated by screening an

A/Reference number: 138144; MUID:93149596; PMID:7678926  
 A/Accession: 138144  
 A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA  
 A/Residues: 'MCMR', 504-767, 'V', 769-802, 'GI' <LIN>  
 A/Cross-references: EMBL:X70500; NID:G312815; PIDN:CAA49912.1; PID:G312816

Genetics:  
 Gene: GDB:TTK  
 A/Cross-references: GDB:455142  
 A/Map position: 7p12-7cen  
 C/Superfamily: protein kinase homology  
 C/Keywords: phosphotransferase  
 F:523-791/Domain: protein kinase homology <KIN>

Query Match 60.8%; Score 45; DB 2; Length 857;  
 Best Local Similarity 69.2%; Pred. No. 16;  
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 NSKSTSKRRGKK 15  
 Db 845 SSKSTPEKRRGKK 857

## RESULT 4

SS8672  
 extra sex combs protein - fruit fly (Drosophila melanogaster)  
 C/Species: Drosophila melanogaster  
 C/Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 26-May-2000  
 C/Accession: S58672  
 R/Gutjahr, T.; Frei, E.; Spicer, C.; Baumgartner, S.; White, R.A.H.; Noll, M.  
 EMBO J. 14, 4296-4306, 1995

A/Title: The polycomb-group gene, extra sex combs, encodes a nuclear member of the WD-4

A/Reference number: S58672; MUID:96016202; PMID:7556071  
 A/Accession: S58672  
 A/Status: preliminary; not compared with conceptual translation  
 A/Molecule type: DNA

A/Residues: 1-425 <GUT>  
 A/Cross-references: GB:L41867; NID:G1050996; PIDN:AAA86427.1; PID:G1050997  
 C/Genetics:  
 A/Status: preliminary  
 A/Accession: S58672

A/Map position: 3  
 A/Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.el  
 A/Accession: G88448  
 A/Status: preliminary  
 A/Accession: G88448

Query Match 59.5%; Score 44; DB 2; Length 425;  
 Best Local Similarity 60.0%; Pred. No. 13;  
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 KONSSTSKRRGKK 15  
 Db 36 KSPSSSTRSKRRGR 50

## RESULT 5

T34103  
 hypothetical protein C17G10.6 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

Accession: T34103  
 R/Johnson, D.  
 submitted to the EMBL Data Library, June 1995  
 A/Description: The sequence of C. elegans cosmid C17G10.  
 A/Reference number: Z21476  
 A/Accession: T34103

A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-683 <JOH>  
 A/Cross-references: EMBL:U28739; PIDN:AB93453.1; GSPDB:GN00020; CESP:C17G10.6

A/Experimental source: strain Bristol NZ; clone C17G10  
 C/Genetics:  
 A/Map position: 2  
 A/Accession: C17G10.6

Query Match 58.1%; Score 43; DB 2; Length 683;  
 Best Local Similarity 53.3%; Pred. No. 28;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 KONSSTSKRRGKK 15  
 Db 557 KKKSKKNNRKRGRK 571

## RESULT 6

G88448  
 protein C45G9.4 [imported] - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001

Accession: G88448  
 R/Anonymous, The C. elegans Sequencing Consortium.  
 Science 285, 2012-2018, 1998  
 A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo

A/Reference number: A75000; MUID:99069613; PMID:9851916  
 A/Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.el  
 A/Accession: G88448  
 A/Status: preliminary  
 A/Accession: G88448

A/Molecule type: DNA  
 A/Residues: 1-292 <STO>  
 A/Cross-references: GB:chr\_III; PIDN:AAA62556.1; PID:G687878; GSPDB:GN00021; CESP:C45G9

Query Match 56.8%; Score 42; DB 2; Length 292;

Best Local Similarity 57.1%; Pred. No. 20;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KNSKTTSKRRGKK 15  
Db 30 RRSKTTSKRRGKK 43

RESULT 7  
E88448  
Protein C45G9.9 [imported] - Caenorhabditis elegans

C/Species: Caenorhabditis elegans  
C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C/Accession: E88448  
R/Anonymous: The C. elegans Sequencing Consortium.  
S/Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A/Reference number: A75000; MUID:99069613; PMID:9851916  
A/Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/projects/C\_elegans/  
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
Accession: E88448  
Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-294 <STO>  
A/Cross-references: GB:chr\_III; PIDN:AAA62552.1; PID:9687874; GSPDB:GN00021; CESP:C45G9.

C/Genetics:  
A/Gene: C45G9.9  
A/Map position: 3

Query Match 56.8%; Score 42; DB 2; Length 294;  
Best Local Similarity 57.1%; Pred. No. 20;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KNSKTTSKRRGKK 15  
Db 30 RRSKTTSKRRGKK 43

RESULT 8

regulatory protein PHO4 - yeast (Saccharomyces cerevisiae)  
N/Alternate names: phosphate system positive regulatory protein; protein R007; protein Y

C/Species: Saccharomyces cerevisiae  
C/Date: 02-Sep-1995 #sequence\_revision 12-Apr-1996 #text\_change 17-Mar-1999  
C/Accession: S56289; A23482; S62245; S63839  
R/Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanna, S.I.; Sasana  
submitted to the EMBL Data Library, May 1995  
Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces ce

Accession number: S56289

A/Molecule type: DNA

A/Residues: 1-312 <MUR>

A/Cross-references: EMBL:D50617; NID:G836685; PID:d1009914; PID:G836789; MIPS:YFR034C

R/Regelin, M.; De Wilde, M.; Hliger, F.

Nucleic Acids Res. 14, 3059-3073, 1986

A/Title: Isolation, physical characterization and expression analysis of the Saccharomy

A/Reference number: A23482; MUID:86176785; PMID:3008105

A/Accession: A23482

A/Molecule type: DNA

A/Residues: 1-289; 'RP', '292', 'RPPWRRPACTSVT', '302', 'S', '304', 'T' <LEG>

A/Cross-references: EMBL:X03719

R/Murakami, Y.

submitted to the EMBL Data Library, December 1994

A/Reference number: S62230

A/Accession: S62230

A/Molecule type: DNA

A/Residues: 1-312 <MUR>

A/Cross-references: EMBL:D44602; NID:G893419; PID:d1008636; PID:G893429

R/Eki, T.; Naitou, M.; Hagiwara, H.; Abe, M.; Ozawa, M.; Sasanna, S.I.; Sasanna, M.; T

Yeast 12, 177-190, 1996

A/Title: Fifteen open reading frames in a 30.8 kb region of the right arm of chromosome

A/Reference number: S63830; MUID:96287654; PMID:8686381

A/Accession: S63830

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-312 <EKT>

A/Cross-references: EMBL:D44602; NID:G893419; PID:d1008636; PID:G893429

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1995

C/Genetics:

A/Gene: SGD:PHO4

A/Cross-references: SGD:S0001930; MIPS:YFR034C

A/Map position: 6R

Query Match 56.8%; Score 42; DB 2; Length 312;  
Best Local Similarity 64.3%; Pred. No. 21;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KNSKTTSKRRGKK 14  
Db 148 KNSSPYLNKRRGKK 161

RESULT 9

hypothetical protein C43F9.6 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C/Accession: T19910

R/Mortimore, B.

submitted to the EMBL Data Library, November 1996

A/Reference number: Z19195

A/Accession: T19910

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-374 <WIL>

A/Cross-references: EMBL:Z82262; PIDN:CA05149.1; GSPDB:GN00022; CESP:C43F9.6

A/Experimental source: clone C43F9

C/Genetics:

A/Gene: CESP:C43F9.6

A/Map position: 4

A/Introns: 22/1; 82/2; 104/1; 151/1; 193/1; 235/2; 292/3; 335/3

Query Match 56.8%; Score 42; DB 2; Length 374;  
Best Local Similarity 69.2%; Pred. No. 25;  
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KNSKTTSKRRGKK 13  
Db 20 KNSKTTAKSRVG 32

RESULT 10

S64750

probable ATP-dependent RNA helicase DR1 - yeast (Saccharomyces cerevisiae)

N/Alternate names: protein L1345; protein YL008w

C/Species: Saccharomyces cerevisiae

C/Date: 01-Aug-1995 #sequence\_revision 24-May-1996 #text\_change 02-Feb-2001

C/Accession: S64750; S31248; S70567

R/Mioega, T.; Zimmermann, F.K.

submitted to the Protein Sequence Database, May 1996

A/Reference number: S64743

A/Accession: S64750

A/Molecule type: DNA

A/Residues: 1-752 <MIO>

A/Cross-references: EMBL:Z73113; NID:G1360170; PIDN:CAA97452.1; PID:E245447; PID:G13601,

A/Experimental source: strain S288C

R/Ripmaster, T.L.; Vaughn, G.P.; Woolford Jr., J.L.

Proc. Natl. Acad. Sci. U.S.A. 89, 1131-1135, 1992

A/Title: A putative ATP-dependent RNA helicase involved in Saccharomyces cerevisiae rib

A/Reference number: S31248; MUID:93087480; PMID:11454790

A/Accession: S31248

A/Molecule type: DNA

A/Residues: 'MKR', '34', 'SRRLRRSGRVITR', '51', 'RLVR', '56', 'ITSMRMFWRTWT', '69-752 <RIP>

A/Cross-references: EMBL:L00683; NID:G171655; PIDN:AAA4666.1; PID:G171656

R/Mioega, T.; Zimmermann, F.K.

Yeast 12, 693-708, 1996

A/Title: Sequence analysis of the GEN12 region of *Saccharomyces cerevisiae* on a 43.7 kb  
e conductance regulator protein CTR.  
A/Reference number: S70557; MUID:96405918; PMID:8810043  
A/Accession: S70567  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-752 <MW>  
A/Cross-references: EMBL:X91488; NID:G1495203; PIDN:CA62783.1; PID:aj99012; PID:G149521  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1995  
C/Genetics:  
A/Map position: 12L  
A/Cross-references: SGD:S0003931; MIPS:YLL008W  
A/Map position: 12L  
C/Keywords: ATP; nucleotide binding; P-loop  
F/275-282/Region: nucleotide-binding motif A (P-loop)  
F/381-386/Region: nucleotide-binding motif B  
F/385-388/Region: DEAD motif

Query Match 56.8%; Score 42; DB 2; Length 752;  
Best Local Similarity 57.1%; Pred. NO. 44;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 KONSSTTSKRGRK 14  
| : ||| : |||  
Db 34 KVEAKTTTKRGRK 47

RESULT 11  
JAC6032  
lactococpin (EC 3.4.21.96) precursor [similarity] - *Lactobacillus delbrueckii* subsp. bulg  
N/Alternate names: cell envelope-associated serine proteinase prtp  
C/Species: *Lactobacillus delbrueckii* subsp. bulgarius  
C/Date: 31-Dec-1996 #sequence\_revision 31-Dec-1996 #text\_change 21-Jul-2000  
C/Accession: J06032  
R/Gilbert, C.; Altan, D.; Blanc, B.; Portaller, R.; Germond, J.E.; Lapierre, L.; Moller, J.  
J. Bacteriol. 178, 3059-3065, 1996  
A/Title: A new cell surface proteinase: Sequencing and analysis of the prtB gene from *La*  
A/Reference number: J06032; MUID:96236017; PMID:8655480  
A/Accession: J06032  
A/Molecule type: DNA  
A/Residues: 1-1946 <GIL>  
A/Cross-references: GB:L48487; NID:G6013471; PIDN:AAC41529.1; PID:G1381114  
A/Experimental source: NCD01489  
A/Note: neither the complete nucleic acid sequence nor the complete translation are show  
C/Genetics:  
A/Map position: prtB  
C/Superfamily: lactococpin; subtilisin homology  
C/Keywords: hydrolase; serine proteinase  
F/213-631/Domain: subtilisin homology #status atypical <SRT>

Query Match 56.8%; Score 42; DB 2; Length 1946;  
Best Local Similarity 60.0%; Pred. NO. 98;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 KONSSTTSKRGRK 15  
| : ||| : |||  
Db 1789 KKTDSKTTSKSARK 1803

RESULT 12  
G96695  
hypothetical protein F5A8.9 [imported] - *Arabidopsis thaliana*  
C/Species: *Arabidopsis thaliana* (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C/Accession: G96695  
R/Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chih, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huilar, L.  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A/Reference number: A66141; MUID:21016719; PMID:11130712  
A/Accession: G96695  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-453 <STO>  
A/Cross-references: GB:AE005173; NID:G4204282; PIDN:AD10663.1; GSPDB:GN00141  
C/Genetics:  
A/Map position: 1

Query Match 55.4%; Score 41; DB 2; Length 453;  
Best Local Similarity 53.3%; Pred. NO. 42;  
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 KONSSTTSKRGRK 15  
| : ||| : |||  
Db 164 KENSGVAESRGRK 178

RESULT 13  
T34308  
hypothetical protein F55C12.5 - *Caenorhabditis elegans*  
C/Species: *Caenorhabditis elegans*  
C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C/Accession: T34308  
R/Natrelle, P.  
Submitted to the EMBL Data Library, November 1995  
A/Description: The sequence of C. elegans cosmid F55C12.  
A/Reference number: Z21503  
A/Accession: T34308  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-895 <LAT>  
A/Cross-references: EMBL:U41107; PIDN:AAC71164.1; GSPDB:GN00020; CESP:F55C12.5  
A/Experimental source: strain Bristol N2; clone F55C12  
C/Genetics:  
A/Map position: F55C12.5  
A/Map position: 2  
A/Intons: 6/3; 25/3; 66/3; 240/3; 481/2; 535/3; 687/3; 791/3; 827/2

Query Match 55.4%; Score 41; DB 2; Length 895;  
Best Local Similarity 53.3%; Pred. NO. 75;  
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 KONSSTTSKRGRK 15  
| : ||| : |||  
Db 752 KNSSEGTTERGRK 766

RESULT 14  
S43434  
histone H1, testicular - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 20-Jun-2000  
C/Accession: S43434; S33226  
R/Draabent, B.; Bode, C.; Doenecke, D.  
Biochim. Biophys. Acta 1216, 311-313, 1993  
A/Title: Structure and expression of the mouse testicular H1 histone gene (H1t).  
A/Reference number: S43434; MUID:94060108; PMID:8241275  
A/Accession: S43434  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-208 <DBA>  
A/Cross-references: EMBL:X72805; NID:G1934959; PIDN:CA51325.1; PID:G297754  
A/Note: the authors did not translate the codon for residue 1  
C/Superfamily: histone H1  
C/Keywords: DNA binding; nucleus

Query Match 54.1%; Score 40; DB 2; Length 208;

Best Local Similarity 80.0%; Pred. No. 32;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 KTTSKRGRKK 15  
|:|||||  
Db 19 KPSSKRGRKK 28

## RESULT 15

C45345

Vif protein - caprine arthritis-encephalitis virus (strain CO)

N:Alternate names: orf-Q protein; sor protein

C:Species: caprine arthritis-encephalitis virus, CAEV

C&gt;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Jul-1999

C:Accession: C45345

R:Salazar, M.; Querat, G.; Konings, D.A.M.; Vigne, R.; Clements, J.E.  
Virology 179, 347-364, 1990A>Title: Nucleotide sequence and transcriptional analysis of molecular clones of CAEV wt  
A:Reference number: A45345; MUID:91021037; PMID:2171210

Molecule type: mRNA

Residues: 1-229 &lt;SAL&gt;

A:Cross-references: GB:M33677; NID:G323294; PIDN:AAA91827.1; PID:G323297

C:Gene: vif

C:Superfamily: Viena virus vif protein

## Query Match

Best Local Similarity 54.1%; Score 40; DB 1; Length 229;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 QNSSKTSKRGRKK 15  
|:|||||  
Db 2 QNSSRHOQKKRKK 15

Search completed: May 1, 2003, 20:38:42  
Job time : 3.89958 sec

4

GenCore version 5.1.4 p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:13:36 ; Search time 0.753138 Seconds

(without alignments)  
826.070 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_407\_421

Perfect score: 74

Sequence: 1 KONSSTTSKRKCK 15

Scoring table:

BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

al number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	62.2	216	1 GCH1_DEIRA	G9RYB4 deinoxococcus
2	45	60.8	841	1 TTK_HUMAN	P33981 homo sapien
3	42	56.8	292	1 VO14_CAEEL	O09505 caenorhabdi
4	42	56.8	294	1 VO19_CAEEL	O09507 caenorhabdi
5	42	56.8	312	1 PHO4_YEAST	P07270 saccharomyc
6	42	56.8	752	1 DRS1_YEAST	P32892 saccharomyc
7	40	54.1	207	1 H1T_MOUSE	Q07133 mus musculu
8	40	54.1	229	1 V1F_CAEVC	P33462 caprine art
9	40	54.1	298	1 YN05_YEAST	P33843 saccharomyc
10	40	54.1	1233	1 YF16_YEAST	P35662 bos taurus
11	39	52.7	667	1 UVRB_ECOLI	P07025 escherichia
12	39	52.7	673	1 CENC_MOUSE	P49452 mus musculu
13	39	52.7	905	1 IF41_YEAST	P08095 saccharomyc
14	39	52.7	952	1 TOP2_SCHPO	P20054 dictyostell
15	39	52.7	1485	1 PYR1_DICDI	P55200 mus musculu
16	39	52.7	2185	1 HRX_MOUSE	O10251 schizosacch
17	39	52.7	3866	1 IF2P_SCHPO	P35067 acropora fo
18	38.5	52.0	1079	1 H2B_ACRRO	P48530 xenopus lae
19	38	51.4	125	1 H2B_ACRRO	P48531 xenopus lae
20	38	51.4	142	1 PTA1_XENLA	P53551 saccharomyc
21	38	51.4	142	1 PTA2_XENLA	P53551 saccharomyc
22	38	51.4	258	1 H1_YEAST	P14997 polyomaviru
23	38	51.4	351	1 COA2_POVBA	O43114 schizosacch
24	38	51.4	351	1 COA2_POVBA	O43114 schizosacch
25	38	51.4	455	1 ORCS_SCHPO	P75034 mycoplasma
26	38	51.4	793	1 YPO6_MYCPN	P75034 mycoplasma
27	38	51.4	814	1 YBS2_MYCPN	P07799 schizosacch
28	38	51.4	814	1 TOP1_SCHPO	P05878 human immun
29	38	51.4	856	1 ENV_HV1SC	P40701 mus musculu
30	38	51.4	914	1 IF42_YEAST	P48725 mus musculu
31	38	51.4	1711	1 CHD1_MOUSE	P05987 dictyostell
32	38	51.4	1920	1 PCNT_MOUSE	
33	37	50.0	327	1 KAPR_DICDI	

34	37	50.0	339	1 RL29_SPICI	O31163 epitelasma
35	37	50.0	376	1 GUNK_FUSOX	P45659 fusarium ox
36	37	50.0	433	1 BLT2_CAEEL	O10655 caenorhabdi
37	37	50.0	471	1 RB97_DROME	O02976 drosophila
38	37	50.0	482	1 Y138_MENJA	O57602 methanococ
39	37	50.0	561	1 SR72_SCHPO	O59787 schizosacch
40	37	50.0	882	1 RA50_PYRFU	P58301 pyrococcus
41	37	50.0	892	1 ATX7_HUMAN	O15265 homo sapien
42	37	50.0	914	1 PBPA_BACSV	P39793 bacillus su
43	37	50.0	939	1 XPC_HUMAN	O01831 homo sapien
44	37	50.0	1093	1 SM14_YEAST	P25102 saccharomyc
45	37	50.0	1123	1 RBM6_HUMAN	P78332 homo sapien

## ALIGNMENTS

RESULT 1  
ID GCH1\_DEIRA STANDARD; PRT; 216 AA.

AC G9RYB4;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE GTP cyclohydrolase I (EC 3.5.4.16) (GTP-CH-I).  
GN FOLE OR DR0036.  
OS Deinoxococcus radiodurans.  
OC Bacteria; Thermus/Deinoxococcus group; Deinococci; Deinococcales;  
OC Deinococcales; Deinococcus.  
OX NCBI\_TaxID=1299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RI;  
RX MEDLINE=2003696; PubMed=10567266;  
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
RA Vamathevan J.D., Lam P., McDonald L., Utterback T., Zaleski C.,  
RA Makarova K.S., Aravind L., Daly M.J., Minkov K.W., Fleischmann R.D.,  
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
RA Fraser C.M.;  
RT "Genome sequence of the radioresistant bacterium Deinococcus  
radiodurans R1.";  
RL Science 286:1571-1577 (1999).  
CC -1- CATALYTIC ACTIVITY: GTP + 2 H(2)O = formate + 2-amino-4-hydroxy-6-  
(erythro-1,2,3-trihydroxypropyl)idihydropteridine triphosphate.  
CC -1- PARTWAY: Tetrahydrofolate biosynthesis; first step.  
CC -1- SUBUNIT: HOMOPOLYMER (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE GTP CYCLOHYDROLASE I FAMILY.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
CC -----  
DR EMBL; AE001867; AAF09628.1; -  
DR HSSP; P27511; IGTP.  
DR TIGR; DR0036;  
DR InterPro; IPR001474; GTP\_cyclohydrol.  
DR Pfam; PF01227; GTP\_cyclohydrol; 1.  
DR ProDom; PD003330; GTP\_cyclohydrol; 1.  
DR TIGRFAMs; TIGR00063; FOLE; 1.  
DR PROSITE; PS00859; GTP\_CYCLOHYDROL\_1; 1.  
DR PROSITE; PS00860; GTP\_CYCLOHYDROL\_2; 1.  
KW One-carbon metabolism; Hydrolyase; Complete proteome.  
FT DISULFID 102 173  
FT SEQUENCE 216 AA; 23799 MW; 149F5281A60D4C4C CRC64;

Query Match. 62.2%; Score 46; DB 1; Length 216;  
Best Local Similarity 76.9%; Pred. No. 0.88;

Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KNSKSTTSRRG 13  
 |||||  
 DB 181 KNSSTTSAMRG 193

## RESULT 2

TTK\_HUMAN STANDARD; PRT; 841 AA.

AC P33981;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Dual specificity protein kinase TTK (EC 2.7.1.-) (PYT).  
 GN TTK OR MPS1L.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NCBI\_TaxID=9606;

[1] SEQUENCE FROM N.A.

RA MEDLINE=92348472; PubMed=1639825;  
 RA Mills G.B., Schmandt R., McGill M., Amendola A., Hill M.,  
 RA Jacobs K., May C., Rodricks A.-M., Campbell S., Hogg D.,  
 RT "Expression of TTK, a novel human protein kinase, is associated with  
 cell proliferation.";  
 RL J. Biol. Chem. 267:16000-16006 (1992).

RL [2] SEQUENCE OF 509-776 FROM N.A.

RA MEDLINE=92065863; PubMed=1956325;  
 RA Linberg R.A., Hunter T.;

RL Unpublished results, cited by:

RL Hanks S.K., Quinn A.M.;

RL Meth. Enzymol. 200:38-62 (1991).

CC -1- FUNCTION: PHOSPHORYLATES PROTEINS ON SERINE, THREONINE, AND

CC -1- TISSUE SPECIFICITY: PRESENT IN RAPIDLY PROLIFERATING CELL LINES.

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: M86699; AAA61239.1; -

PIR: S27971; S27971.

DR HSSP: 000534; 1B18.

DR Genew: MGN:12401; TTK.

DR MIM: 604092; -

DR InterPro: IPR000719; Euk.pkinase.

DR InterPro: IPR002290; Ser\_thr\_pkinase.

DR Pfam: P000069; pkinase; 1.

DR ProDom: PD000001; Euk.pkinase; 1.

DR SMART: SM00220; S\_TKC; 1.

DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.

DR PROSITE: PS00108; PROTEIN KINASE ST; 1.

DR PROSITE: PS00111; PROTEIN KINASE DOM; 1.

KW Transferase; Serine/threonine-protein kinase; ATP-binding;

KW Tyrosine-protein kinase.

FT DOMAIN 509 775 PROTEIN KINASE.

FT NP BIND 515 523 ATP (BY SIMILARITY).

FT BINDING 537 537 ATP (BY SIMILARITY).

FT ACT SITE 631 631 BY SIMILARITY.

FT CONFIDENT 752 752 L -> V (IN REF. 2).

FT SEQUENCE 841 AA; 95279 MW; ACDECA2CD15EF8C CRC64;

Query Match 60.8%; Score 45; DB 1; Length 841;

Best Local Similarity 69.2%; Pred. No. 5.2;

Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 NSSKTSRRGK 15  
 :|||:|  
 DB 829 SSSKTFKRRGK 841

## RESULT 3

Y014\_CAEEL STANDARD; PRT; 292 AA.

ID Y014\_CAEEL  
 AC Q09505;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE Hypothetical 32.4 kDa protein C45G9.4 in chromosome III.  
 GN C45G9.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 NCBI\_TaxID=6239;

[1] SEQUENCE FROM N.A.

RA STRAIN=Bristol N2;

RA Bentley D., Waterston R.;

RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: U21323; AAA62556.1; -

DR WormRep: C45G9.4; CRO1855.

DR Hypothetical protein.

FT DOMAIN 9 84 LYS-RICH.

SO SEQUENCE 292 AA; 32454 MW; 589FCD425363A988 CRC64;

Query Match 56.8%; Score 42; DB 1; Length 292;

Best Local Similarity 57.1%; Pred. No. 5.6;

Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 QNSKTSRRGK 15  
 :|||:|  
 DB 30 RRSKTSRRGK 43

## RESULT 4

Y019\_CAEEL STANDARD; PRT; 294 AA.

ID Y019\_CAEEL  
 AC Q09507;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE Hypothetical 32.7 kDa protein C45G9.9 in chromosome III.  
 GN C45G9.9.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 NCBI\_TaxID=6239;

[1] SEQUENCE FROM N.A.

RA STRAIN=Bristol N2;

RA Bentley D., Waterston R.;

RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>



or send an email to [license@sib.ch](mailto:license@sib.ch).

CC -----

DR EMBL; U21323; AAA62552.1; -

DR WormPep; C4569.9; CE01860.

KW Hypothetical protein.

FT DOMAIN

SO SEQUENCE 294 AA; 32738 MW; 856B595C1FCC448 CRC64;

LYS-RICH.

Query Match 56.8%; Score 42; DB 1; Length 294;

Best Local Similarity 57.1%; Pred. No. 5.7;

Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 QNSKSTSKRRGK 15

DB 30 RRSKTSKRRGK 43

RESULT 5

PH04\_YEAST STANDARD; PRT; 312 AA.

P07270:

01-APR-1988 (Rel. 07, Created)

01-FEB-1991 (Rel. 17, Last sequence update)

15-JUN-2002 (Rel. 41, Last annotation update)

DE Phosphate system positive regulatory protein PHO4.

GN PH04 OR YPR034C.

OS Saccharomyces cerevisiae (Baker's yeast).

CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI\_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=86176785; PubMed=3008105;

RA Legrain M., de Wilde M., Hilger F.;

RT "Isolation, physical characterization and expression analysis of the

RT Saccharomyces cerevisiae positive regulatory gene PHO4.";

RL Nucleic Acids Res. 14:3059-3073(1986).

RL [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=89364688; PubMed=2505053;

RA Yoshida K., Kuromitsu Z., Ogawa N., Oshima Y.;

RT "Mode of expression of the positive regulatory genes PHO2 and PHO4 of

RT the phosphatase regulon in Saccharomyces cerevisiae.";

RL Mol. Genet. 217:31-39(1989).

RL [3]

RP SEQUENCE FROM N.A.

RX STRAIN=S288C / AB972;

RA MEDLINE=95400292; PubMed=7670463;

RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,

RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,

RT "Analysis of the nucleotide sequence of chromosome VI from

RT Saccharomyces cerevisiae.";

RL Nat. Genet. 10:261-268(1995).

RL [4]

RP SEQUENCE FROM N.A.

RX STRAIN=S288C / AB972;

RA MEDLINE=96287654; PubMed=8686381;

RA Eki T., Naitou M., Hagiwara H., Abe M., Ozawa M., Sasanuma S.-I.,

RA Sasanuma M., Tsuchiya Y., Shibata T., Watanabe K., Ono A.,

RA Yamazaki M.-A., Tashiro H., Hanaoka F., Murakami Y.;

RT "Fifteen open reading frames in a 30.8 kb region of the right arm of

RT chromosome VI from Saccharomyces cerevisiae.";

RL Yeast 12:177-190(1996).

RL [5]

RP DOMAINS.

RX MEDLINE=90220608; PubMed=2183025;

RA Ogawa N., Oshima Y.;

RT "Functional domains of a positive regulatory protein, PHO4, for

RT transcriptional control of the phosphatase regulon in Saccharomyces

RT cerevisiae.";

RL Mol. Cell. Biol. 10:2224-2236(1990).

RL [6]

RP HELIX-LOOP-HELIX MOTIF.

RX MEDLINE=91021495; PubMed=2220078;

RA Berben G., Legrain M., Gillionet V., Hilger F.;

RT "The yeast regulatory gene PHO4 encodes a helix-loop-helix motif.";

RL Yeast 6:451-454(1990).

RL [7]

RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 251-312.

RX MEDLINE=97447573; PubMed=9303313;

RA Shimizu T., Tomoto A., Ihara K., Shimizu M., Kyogoku Y., Ogawa N.,

RA Oshima Y., Hoshikawa T.,

RT "Crystal structure of PHO4 bHLH domain-DNA complex: flanking base

RT recognition.";

RL EMBL J. 16:4689-4697(1997).

CC -1- FUNCTION: THIS ACTIVATOR REGULATES THE EXPRESSION OF REPRESSIBLE

CC PHOSPHATASE AT THE TRANSCRIPTIONAL LEVEL. PHO4 IS PRESUMED TO

CC INTERACT WITH THE UAS (UPSTREAM ACTIVATING SEQUENCE) OF SEVERAL

CC PHOSPHATASE ENCODING PHO GENES.

CC -1- SUBUNIT: BINDS DNA AS A DIMER OR HIGHER COMPLEX.

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- PTM: PHOSPHORYLATED BY PHO85 WHEN INORGANIC PHOSPHATE IS PRESENT.

CC THIS PREVENTS PHO4 FROM ACTIVATING THE STRUCTURAL PHO GENES.

CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF

CC TRANSCRIPTION FACTORS.

CC -1- SIMILARITY: TWO REGIONS OF STRONG HOMOLOGY TO PHO2 ARE FOUND.

CC ALSO SIMILAR TO N. CRASSA NUC1.

CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN IN THE C-TERMINUS

CC DUE TO A FRAMESHIFT.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>

CC or send an email to [license@sib.ch](mailto:license@sib.ch)).

CC -----

DR EMBL; X03719; CAA27345.1; ALT\_FRAME.

DR EMBL; D50617; BAA09273.1; -.

DR PIR; A23482; A23482.

DR PDB; 1A0A; 18-MAR-98.

DR TRNSPAC; T00690; -.

DR SGD; S0001930; PHO4.

DR InterPro; IPR001092; HLH\_basic.

DR Pfam; PF00010; HLH; 1.

DR SMART; SM00353; HLH; 1.

DR PROSITE; PS00038; HLH; 1.

DR PROSITE; PSS0888; HLH; 2; 1.

KW Transcription regulation; Activator; DNA-binding; Phosphorylation;

KW Nuclear protein; 3D-structure.

FT DOMAIN 1 109

FT DOMAIN 203 227

FT DNA BIND 251 307

FT DOMAIN 2 307

FT CONFLICT 269 269

FT CONFLICT 310 310

FT CONFLICT 310 310

SO SEQUENCE 312 AA; 34072 MW; 6DB926A98A09241 CRC64;

Query Match 56.8%; Score 42; DB 1; Length 312;

Best Local Similarity 64.3%; Pred. No. 6;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KNSKSTSKRRGK 14

DB 148 KNSKSTSKRRGK 161

RESULT 6

DRS1\_YEAST STANDARD; PRT; 752 AA.

AC P32882;

DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Probable ATP-dependent RNA helicase DRS1.  
 GN DRS1 OR YL008W OR L1345.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 ON NCBI\_TaxID=4932;  
 RX MEDLINE=93087480; PubMed=1454790;  
 RP MEDLINE=93087480; PubMed=1454790;  
 RA Rpmaster T.L., Vaughn G.P., Woolford J.L. Jr.;  
 RT "A putative ATP-dependent RNA helicase involved in Saccharomyces  
 cerevisiae ribosome assembly";  
 RT Proc. Natl. Acad. Sci. U.S.A. 89:11131-11135(1992).  
 RN [2]  
 RL SEQUENCE FROM N.A.  
 RC STRAIN=5288c / FY23;  
 RX MEDLINE=96405918; PubMed=8810043;  
 RA Moosge T., Zimmermann P.K.;  
 RT "Sequence analysis of the CEN12 region of Saccharomyces cerevisiae on  
 a 43.7 kb fragment of chromosome XII including an open reading frame  
 homologous to the human cystic fibrosis transmembrane conductance  
 regulator protein CFTR";  
 RL Yeast 12:693-708(1996).  
 CC -1- FUNCTION: PROBABLE HELICASE INVOLVED IN RIBOSOME ASSEMBLY.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L00683; AAA3466.1; -;  
 DR EMBL: X91488; CAA62783.1; -;  
 DR EMBL: Z73113; CAA97452.1; -;  
 DR PIR: S31248; S31248.  
 DR HSSP: Q58083; IHV8.  
 DR SCD: S0003931; DRS1.  
 DR InterPro: IPR001410; DEAD.  
 DR InterPro: IPR000629; DEAD box.  
 DR InterPro: IPR001650; Helicase\_C.  
 DR Pfam: PF00270; DEAD\_1.  
 DR Pfam: PF00271; Helicase\_C\_1.  
 DR SMART: SM00487; DEXDC\_1.  
 DR SMART: SM00490; HELIC\_C\_1.  
 DR PROSITE: PS00039; DEAD ATP HELICASE; 1.  
 DR ATP-binding; RNA-binding; Helicase; Nuclear protein.  
 FT DOMAIN 170 190 POLY-GU.  
 FT NP\_BIND 275 282 ATP (POTENTIAL).  
 FT SITE 385 388 DEAD BOX.  
 FT CONFLICT 1 MVTGTRKYSNLDPVFTISDSDVPIIDSDDEKVEAKKT  
 FT KKKRGNKKKRVSEGDINDDEVDL -> MTKSRRLRL  
 FT RRSGRVTRTRKRLVGIYSMMFMRTWT (IN REF. 1).  
 SO SEQUENCE 752 AA; 84643 MW; 60747607A6E5E4A8 CRC64;  
 QY Query Match 56.8%; Score 42; DB 1; Length 752;  
 Best Local Similarity 57.1%; Pred. No. 15;  
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 DB 1 KONSSTTSKRKK 14  
 DB 34 KVEAKKTKRKK 47  
 RESULT 7  
 HIT\_MOUSE STANDARD; PRT; 207 AA.  
 ID HIT\_MOUSE  
 AC Q07133;  
 DT 01-OCT-1994 (Rel. 30; Created)  
 DT 01-NOV-1995 (Rel. 32; Last sequence update)

DT 16-OCT-2001 (Rel. 40; Last annotation update)  
 DE Histone H1 (testicular H1 histone).  
 GN HIST OR HIT OR H1P3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;  
 RX MEDLINE=94060108; PubMed=8241275;  
 RP MEDLINE=94060108; PubMed=8241275;  
 RA Drabant B., Bode C., Doenecke D.;  
 RT "Structure and expression of the mouse testicular H1 histone gene  
 (H1t)";  
 RT Biochim. Biophys. Acta 1216:311-313(1993).  
 RN [2]  
 RL SEQUENCE FROM N.A.  
 RC STRAIN=CD-1; TISSUE=Testis;  
 RP STRAIN=CD-1; TISSUE=Testis;  
 RA van Wert J., Wright J., Wolfe S.A., Grimes S.R.;  
 RT Submitted (FEB-1994) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: HISTONES H1 ARE NECESSARY FOR THE CONDENSATION OF  
 CC NUCLEOSOME CHAINS INTO HIGHER ORDER STRUCTURES.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- DEVELOPMENTAL STAGE: THIS HISTONE IS A TESTIS-SPECIFIC H1 VARIANT  
 CC THAT APPEARS DURING MEIOSIS IN SPERMATOGENESIS.  
 CC -1- SIMILARITY: BELONGS TO THE HISTONE H1/H5 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L28753; AAB38417.1; -;  
 DR EMBL: X72805; CAA51325.1; -;  
 DR EMBL: U06232; AAA18359.1; -;  
 DR PIR: S33226; S33226.  
 DR PIR: S43434; S43434.  
 DR HSSP: P08287; IGHC.  
 DR WGD: MG1:107502; H1E3.  
 DR InterPro: IPR001386; Histone\_H1/H5.  
 DR Pfam: PF00538; linker histone; 1.  
 DR ProDom: PD000373; Linkerhist\_N; 1.  
 DR SMART: SM00526; H15; 1.  
 DR Chromosomal protein; Nuclear protein; DNA-binding; Multigene family;  
 KW Acetylation; Spermatogenesis; Testis.  
 FT INIT MET 0 0 BY SIMILARITY.  
 FT MOD RES 1 1 ACETYLATION (BY SIMILARITY).  
 FT DOMAIN 37 110 GLOBULAR.  
 FT CONFLICT 163 163 R -> G (IN REF. 1).  
 SO SEQUENCE 207 AA; 21508 MW; A3C847CA97976C44 CRC64;  
 QY Query Match 54.1%; Score 40; DB 1; Length 207;  
 Best Local Similarity 80.0%; Pred. No. 8; 6;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 DB 6 KTSKRKK 15  
 DB 18 KPSSKRKK 27  
 RESULT 8  
 VIF\_CAEVC STANDARD; PRT; 229 AA.  
 ID VIF\_CAEVC  
 AC P33462;  
 DT 01-FEB-1994 (Rel. 28; Created)  
 DT 01-FEB-1994 (Rel. 28; Last sequence update)  
 DT 01-FEB-1994 (Rel. 28; Last annotation update)  
 DE Varion infectivity factor (Q protein) (SOR protein).  
 GN VIF.  
 OS Caprine arthritis encephalitis virus (strain Corb) (CAEV).

OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 RX NCBI\_TaxID=11661;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91021037; PubMed=2171210;  
 RA Salterelli M., Querat G., Konings D.A.M., Vigne R., Clements J.E.;  
 RT "Nucleotide sequence and transcriptional analysis of molecular clones  
 of CAEV which generate infectious virus";  
 RL Virology 179:347-364(1990).  
 CC -1- FUNCTION: DETERMINES VIRUS INFECTIVITY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M3677; AAA1827.1; -  
 FIR; C45345; C45345.  
 SEQUENCE 229 AA; 28051 MW; BF78AA3CAB23A101 CRC64;

Query Match 54.1%; Score 40; DB 1; Length 229;  
 Best Local Similarity 57.1%; Pred. No. 9.5;  
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 ONSSKTTSKRRGK 15  
 DB 2 ONSSRRHQKKRKK 15

## RESULT 9

YN05\_YEAST STANDARD; PRT; 298 AA.

AC P53843;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 34.5 kDa protein in PIK1-POL2 intergenic region.  
 GN YN1265C OR N0809.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 CC NCBI\_TaxID=4932;  
 RN [1]

SEQUENCE FROM N.A.

STRAIN=S288C / FY1679;  
 MEDLINE=96310631; PubMed=8740425;

Sen-Gupta M., Lyck R., Fleig U., Niedenthal R.K., Hegemann J.H.;  
 RT "The sequence of a 24,152 bp segment from the left arm of chromosome  
 XIV from Saccharomyces cerevisiae between the BNI1 and the POL2  
 genes";  
 RL Yeast 12:505-514(1996).  
 CC -1- SIMILARITY: TO HUMAN KIAA0174 AND S. POMBE SPOC285.02C.

CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X92494; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; Z71542; CA96172.1; -  
 DR SGD; S0005209; YNL265C.  
 DR InterPro; IPR005061; DUF292.  
 DR Pfam; PF03398; DUF292; 1.  
 KM Hypothetical protein.  
 FT DOMAIN 190 196 POLY-SER.  
 SQ SEQUENCE 298 AA; 34495 MW; F16A94BFC7716A1A CRC64;

Query Match 54.1%; Score 40; DB 1; Length 298;  
 Best Local Similarity 53.3%; Pred. No. 12;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGK 15  
 DB 264 KKSSEKTTKRRKK 278

## RESULT 10

YF16\_YEAST STANDARD; PRT; 1233 AA.

AC P43597;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE Hypothetical 137.7 kDa protein in UGS1-FAB1 intergenic region.  
 GN YFR016C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 CC NCBI\_TaxID=4932;  
 RN [1]

SEQUENCE FROM N.A.  
 STRAIN=S288C / AB972;  
 RX MEDLINE=95400292; PubMed=7670463;  
 RA Murakami Y., Naitou M., Hagihara H., Shibata T., Ozawa M.,  
 RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,  
 RA Yamazaki M., Tashiro H., Eki T.;  
 RT "Analysis of the nucleotide sequence of chromosome VI from  
 RT Saccharomyces cerevisiae";  
 RL Nat. Genet. 10:261-268(1995).  
 CC -1- SIMILARITY: SOME, TO MAMMALIAN NEUROFILAMENT TRIPLET M PROTEIN.

CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; D50617; BAA09255.1; -  
 DR SGD; S0001912; YFR016C.  
 KM Hypothetical protein.  
 SQ SEQUENCE 1233 AA; 137697 MW; C8A7CD2C6F0892P6 CRC64;

Query Match 54.1%; Score 40; DB 1; Length 1233;  
 Best Local Similarity 50.0%; Pred. No. 53;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGK 14  
 DB 186 KKKKKKTTTTRRR 199

## RESULT 11

CYL1\_BOVIN STANDARD; PRT; 667 AA.

AC P35662;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Cyclicin I (Multiple-band polypeptide I).  
 GN CYL1 OR CYL.

OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Testis;

RX MEDLINE=93359502; PubMed=8354692;  
 RA Hess H., Held H., Franke W.W.;  
 RT Molecular characterization of mammalian cyclin, a basic protein of  
 the sperm head cytoskeleton.";  
 RL J. Cell Biol. 122:1043-1052(1993).  
 CC -1- FUNCTION: POSSIBLE ARCHITECTURAL ROLE DURING SPERMATOGENESIS. MAY  
 BE INVOLVED IN SPERMATID DIFFERENTIATION.  
 CC -1- SUBCELLULAR LOCATION: CALYX; SPERM HEAD CYTOSKELETAL STRUCTURE.  
 CC -1- TISSUE SPECIFICITY: TESTIS.  
 CC -1- DEVELOPMENTAL STAGE: SPECIFIC TO LATE SPERMATOGENESIS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: Z22779; CA80456.1; -  
 CC PIR: S35913; S35913.  
 CC PIR: A40713; A40713.  
 KW Cytoskeleton; Structural protein; Repeat; Sperm; Spermatogenesis.  
 FT DOMAIN 287 569 9 APPROXIMATE TANDEM REPEATS.  
 FT REPEAT 287 305 1.  
 FT REPEAT 306 337 2.  
 FT REPEAT 338 368 3.  
 FT REPEAT 369 405 4.  
 FT REPEAT 406 442 5.  
 FT REPEAT 443 475 6.  
 FT REPEAT 476 516 7.  
 FT REPEAT 517 547 8.  
 FT REPEAT 548 569 9.  
 FT DOMAIN 617 667 PRO-RICH.  
 FT SEQUENCE 667 AA; 74817 MW; CBP6EA462243D91 CRC64;  
 SQ  
 Query Match 52.7%; Score 39; DB 1; Length 667;  
 Best Local Similarity 53.3%; Fred. No. 42;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 Oy 1 KONSSTTSKRGGK 15  
 Db 596 KMPRRITFKKGGK 610

RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of *Escherichia coli* K-12.";  
 RL Science 277:1453-1474(1997).  
 CC [4]  
 CC SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";  
 RL Nature 409:529-533(2001).  
 CC [5]  
 CC SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RIMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
 RA Kihara S., Shiba T., Hattori M., Shinagawa H.,  
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*  
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 CC [6]  
 CC SEQUENCE OF 168-673 FROM N.A.  
 RP MEDLINE=97061202; PubMed=8905232;  
 RC Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,  
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K.,  
 RA Keshimoko K., Kimura S., Kitagawa M., Makino K., Masuda S.,  
 RA Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y.,  
 RA Nishimoto H., Nishio Y., Saito N., Sempel G., Seki Y., Tagami H.,  
 RA Takemoto K., Wada C., Yamamoto Y., Yano M., Horuchi T.;  
 RT "A 718-kb DNA sequence of the *Escherichia coli* K-12 genome  
 RT corresponding to the 12.7-28.0 min region on the linkage map.";  
 RL DNA Res. 3:137-155(1996).  
 CC [7]  
 CC X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 619-673.  
 RP MEDLINE=20123894; PubMed=10631326;  
 RA Sohi M., Alexandrovich A., Woolenaar G., Visse R., Goosen N.,  
 RA Verne J., Fontecilla-Camps J.C., Champness J., Sanderson M.R.;  
 RT "Crystal structure of *Escherichia coli* UvrB C-terminal domain, and a  
 RT model for UvrB-UVC interaction.";  
 RL FEBS Lett. 465:161-164(2000).  
 CC [8]  
 CC STRUCTURE BY NMR OF 619-673.  
 RP MEDLINE=99297571; PubMed=10371161;  
 RA Alexandrovich A., Sanderson M.R., Woolenaar G.F., Goosen N.,  
 RA Lane A.N.;  
 RT "NMR assignments and secondary structure of the UvrB binding domain of  
 RT UvrB.";  
 RL FEBS Lett. 451:181-185(1999).  
 CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT  
 CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS  
 CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRB STIMULATES  
 CC THE ATPASE ACTIVITY OF UVRB IN THE PRESENCE OF UV-IRRADIATED  
 CC DOUBLE-STRANDED DNA. IT ALSO ENHANCES THE ABILITY OF UVRB TO BIND  
 CC TO UV-IRRADIATED DUPLEX DNA.  
 CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS: UVRB, UVRB AND UVRB.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE UVRB FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 UVR DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X03722; CAA27357.1; -  
 DR EMBL; X03678; CAA27314.1; -  
 DR EMBL; AE000180; AAC73866.1; -  
 DR EMBL; AE000259; AAG55150.1; -  
 DR EMBL; AP002553; BAB34280.1; -  
 DR EMBL; D90716; BAB35437.1; -  
 DR PIR; A23765; BVECB. -  
 DR PDB; 1Q0V; 10-NOV-00.  
 DR ECO2DBASE; C080.0; 6TH EDITION.  
 DR EcGene; EG11062; UVRB.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR InterPro; IPR004807; UVRB.  
 DR InterPro; IPR001943; UVRB/C.  
 DR Pfam; PF00271; helicase\_C; 1.  
 DR Pfam; PF02151; UVR; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00490; HELIC; 1.  
 DR TIGRFAMs; TIGR00631; UVRB; 1.  
 DR PROSITE; PS50151; UVR; 1.  
 DR SOS response; Excision nuclease; DNA repair; ATP-binding;  
 DR 3D-structure; Complete proteome.  
 DR NP\_BIND; 39 46 ATP (POTENTIAL).  
 DR DOMAIN; 633 668 UVR.  
 DR SITE; 630 631 CLEAVAGE.  
 DR CONFLICT; 477 477 H -> R (IN REF. 2).  
 DR SEQUENCE 673 AA; 76226 MW; 2F17204534FDAD7 CRC64;

Query Match 52.7%; Score 39; DB 1; Length 673;  
 Best Local Similarity 61.5%; Pred. No. 42;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 QNSKTTSKRRGK 14  
 Db 601 QNIAKTKAKRGK 613

RESULT 13  
 CENC\_MOUSE STANDARD; PRT; 906 AA.  
 P49452;  
 01-FEB-1996 (Rel. 33, Created)  
 01-FEB-1996 (Rel. 33, Last sequence update)  
 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Centromere protein C (CENP-C) (Centromere autoantigen C).  
 OS CENPC.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Swiss;  
 RX MEDLINE=85048391; PubMed=7959789;  
 RA McKay S., Thomson E., Cooke H.;  
 RT "Sequence homologies and linkage group conservation of the human and  
 RT mouse Cencp genes";  
 RL Genomics 22:36-40(1994).  
 CC -1- FUNCTION: COMPONENT OF THE INNER KINETOCORE PLATE. REQUIRED FOR  
 CC NORMAL KINETOCORE ASSEMBLY.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U03113; AAC04314.1; -  
 DR MGD; MGI.99700; Cencp.  
 DR Chromosomal protein; Nuclear protein; DNA-binding; Phosphorylation;  
 DR Centromere.  
 DR SEQUENCE 906 AA; 102225 MW; 43D529B202E9D71E CRC64;

Query Match 52.7%; Score 39; DB 1; Length 906;  
 Best Local Similarity 46.7%; Pred. No. 57;  
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 KNSKTTSKRRGK 15  
 Db 451 KTNSTQTKERSGK 465

RESULT 14  
 IF41 YEAST  
 ID\_IF41 YEAST STANDARD; PRT; 952 AA.  
 AC P3935;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Eukaryotic initiation factor 4F subunit p150 (eIF-4F) (mRNA cap-  
 DE binding protein complex subunit p150).  
 GN TIF4631 OR YGR162W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycaceae;  
 NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C;  
 RX MEDLINE=97435481; PubMed=9290212;  
 RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;  
 RA Woolford J.L., Trachsel H., Sonnenberg N.;  
 RT "TIF4631 and TIF4632: two yeast genes encoding the  
 RT high-molecular-weight subunits of the cap-binding protein complex  
 RT (eukaryotic initiation factor 4F) contain an RNA recognition  
 RT motif-like sequence and carry out an essential function";  
 RL Mol. Cell. Biol. 13:4860-4874(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C;  
 RX MEDLINE=97435481; PubMed=9290212;  
 RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;  
 RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae  
 RT chromosome VII";  
 RL Yeast 13:1077-1090(1997).  
 CC -1- FUNCTION: INTERACTS WITH THE MRNA CAP STRUCTURE, M7GPPX. THIS  
 CC INTERACTION IS REQUIRED FOR EFFICIENT RIBOSOME BINDING TO THE  
 CC MRNA. TIF4631 IS PROBABLY ESSENTIAL WHEN TIF4632 IS MISSING.  
 CC -1- SUBUNIT: THE CAP-BINDING PROTEIN COMPLEX IS COMPOSED OF AT LEAST  
 CC TWO PROTEINS, A 24 kDa (TIF45) AND A 150-200 kDa SUBUNIT  
 CC (TIF4631).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; I16923; AAA02757.1; -  
 DR EMBL; Z72947; CAA97184.1; -  
 DR PIR; A48086; A48086.  
 DR SGD; S0003394; TIF4631.  
 DR InterPro; IPR003890; IF\_eIF4G.  
 DR InterPro; IPR00504; RNA\_rec\_mot.



Sun May 4 10:01:34 2003

us-09-757-982-5\_copy\_407\_421.rsp

Page 9

Search completed: May 1, 2003, 20:36:30  
Job time : 10.7531 secs

---





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:28:47 ; Search time 1.67364 Seconds  
(without alignments)  
1846.697 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_407\_421

Perfect score: 74

Sequence: 1 KONSSTTSKRGRKK 15

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Actual number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: SP\_ARCHAEA:\*  
2: SP\_BACTERIA:\*  
3: SP\_FUNGI:\*  
4: SP\_HUMAN:\*  
5: SP\_INVERTEBRATE:\*  
6: SP\_MAMMAL:\*  
7: SP\_MHC:\*  
8: SP\_ORGANELLE:\*  
9: SP\_PHAGE:\*  
10: SP\_PLANT:\*  
11: SP RODENT:\*  
12: SP\_VIRUS:\*  
13: SP\_VERTEBRATE:\*  
14: SP\_UNCLASSIFIED:\*  
15: SP\_VIRUS:\*  
16: SP\_BACTERIAP:\*  
17: SP\_ARCHAEP:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	100.0	455	4 Q9HC4	Q9HC4 homo sapien
2	69	93.2	454	11 Q9ESL3	Q9ESL3 mus musculi
3	48	64.9	241	11 Q9CRY1	Q9CRY1 mus musculi
4	48	64.9	1117	11 Q88720	Q88720 mus musculi
5	48	64.9	1122	11 Q88388	Q88388 mus musculi
6	46	62.2	530	10 Q9XID7	Q9XID7 arabidopsis
7	45	60.8	857	4 Q9BWS1	Q9BWS1 homo sapien
8	44	59.5	425	5 Q24338	Q24338 arabidopsis
9	43	58.1	124	10 Q9SGW6	Q9SGW6 arabidopsis
10	43	58.1	270	13 Q9PUX7	Q9PUX7 lampyris sp.
11	43	58.1	556	10 Q8RYU7	Q8RYU7 oryza sativ
12	43	58.1	683	5 Q09977	Q09977 caenorhabdi
13	42	56.8	271	3 Q9UVP0	Q9UVP0 candida gla
14	42	56.8	374	5 Q9XUC9	Q9XUC9 caenorhabdi
15	42	56.8	1946	2 Q48545	Q48545 lactobacilli
16	41	55.4	425	5 Q26458	Q26458 drosophila

17	41	55.4	453	10 Q9ZWR9	Q9ZWR9 arabidopsis
18	41	55.4	761	4 Q9ULI9	Q9ULI9 homo sapien
19	41	55.4	876	5 Q95ZS8	Q95ZS8 caenorhabdi
20	41	55.4	925	5 Q95ZS7	Q95ZS7 caenorhabdi
21	40	54.1	560	5 Q96139	Q96139 plasmodium
22	40	54.1	3978	5 Q97236	Q97236 plasmodium
23	39	52.7	82	17 Q8TJ10	Q8TJ10 methanosarc
24	39	52.7	107	10 Q64841	Q64841 mus musculi
25	39	52.7	166	11 Q9DSN1	Q9DSN1 mus musculi
26	39	52.7	241	10 Q9FEJ8	Q9FEJ8 marchantia
27	39	52.7	254	10 Q9RXJ5	Q9RXJ5 arabidopsis
28	39	52.7	273	13 Q9PUX0	Q9PUX0 sarcocentro
29	39	52.7	312	3 Q06859	Q06859 saccharomyc
30	39	52.7	312	3 Q06860	Q06860 saccharomyc
31	39	52.7	319	16 Q8YF97	Q8YF97 bruceella me
32	39	52.7	325	10 Q9SNS9	Q9SNS9 oryza sativ
33	39	52.7	328	5 Q9VAH8	Q9VAH8 drosophila
34	39	52.7	328	5 Q9VBB3	Q9VBB3 drosophila
35	39	52.7	382	12 Q72591	Q72591 beet soil-b
36	39	52.7	401	4 Q9UFC3	Q9UFC3 homo sapien
37	39	52.7	439	4 Q8TCT1	Q8TCT1 homo sapien
38	39	52.7	441	12 Q9TH81	Q9TH81 porcine epi
39	39	52.7	441	12 Q07499	Q07499 porcine epi
40	39	52.7	446	10 Q94CT7	Q94CT7 oryza sativ
41	39	52.7	485	5 Q9N821	Q9N821 trypanosoma
42	39	52.7	546	5 Q22765	Q22765 caenorhabdi
43	39	52.7	548	12 Q8U2C5	Q8U2C5 chelonius in
44	39	52.7	615	5 Q61609	Q61609 acropora ml
45	39	52.7	627	4 Q9H8G3	Q9H8G3 homo sapien

## ALIGNMENTS

## RESULT 1

Q9HC4; PRELIMINARY; PRT; 455 AA.  
ID Q9HC4  
AC Q9HC4; 01-MAR-2001 (T-EMBLrel. 16, Created)  
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)  
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
DE MTKR-beta (Similar to sterile-alpha motif and leucine zipper  
DE containing kinase AZK) (Mixed lineage kinase) (Mixed lineage  
DE kinase-related kinase MKK-beta).  
GN MTKR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gotch I., Adachi M., Nishida E.;  
RT "Identification and Characterization of a Novel MAP Kinase Kinase  
RT kinase, MTKR.";  
RL Submitted (OCT-2000) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=COLON;  
RA Strausberg R.;  
RL Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Acton S.;  
RT "MLK-mixed lineage kinase.";  
RL Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=21950776; PubMed=11836244;  
RA Gross E.A., Callow M.G., Waldbaum L., Thomas S., Ruggieri R.;  
RT "MRK, a Mixed Lineage Kinase-related Molecule That Plays a Role in  
RT gamma-Radiation-Induced Cell Cycle Arrest.";  
RL J. Biol. Chem. 277:13873-13882(2002).  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL: AB049734; BAB16445.1; -  
 DR EMBL: BC001401; AAB01401.1; -  
 DR EMBL: AF325454; AAK1615.1; -  
 DR EMBL: AF480462; AAL85892.1; -  
 DR HSSP: P12931; IFMK.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR004040; STY\_pkinase.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00069; Pkinase.1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR ProDom: PD000001; Euk\_pkinase.1.  
 DR SMART: SM00221; STYKc.1.  
 DR SMART: SM00220; S\_TYKc.1.  
 DR SMART: SM00219; TYRKc.1.  
 DR PROSITE: PS00011; PROTEIN KINASE DOM.1.  
 DR PROSITE: PS00108; PROTEIN KINASE ST.1.  
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 455 AA; 51582 MW; E87DB84A4D58B752 CRC64;

Query Match 100.0%; Score 74; DB 4; Length 455;  
 Best Local Similarity 100.0%; Pred. No. 6,8e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KONSSTTSKRRGKK 15  
 DB 407 KONSSTTSKRRGKK 421

RESULT 2  
 ID Q9ESL3 PRELIMINARY; PRT; 454 AA.  
 AC Q9ESL3;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE MLTK-beta.  
 GN ZAK OR MLTK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21264927; PubMed=11042189;  
 RA Gotch I., Adachi M., Nishida E.;  
 RT "Identification and Characterization of a Novel MAP Kinase Kinase  
 PT J. Biol. Chem. 276:4276-4286(2001).  
 -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL: AB049732; BAB16443.1; -  
 DR HSSP: P12931; IFMK.  
 DR MGD: MGI:1931274; Zak.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR004040; STY\_pkinase.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00069; Pkinase.1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR ProDom: PD000001; Euk\_pkinase.1.  
 DR SMART: SM00221; STYKc.1.  
 DR SMART: SM00220; S\_TYKc.1.  
 DR SMART: SM00219; TYRKc.1.  
 DR PROSITE: PS00011; PROTEIN KINASE DOM.1.  
 DR PROSITE: PS00108; PROTEIN KINASE ST.1.  
 DR ATP-binding; Serine/threonine-protein kinase; Transferase.  
 KW SEQUENCE 454 AA; 51366 MW; 35C2FCOD729D395 CRC64;

Query Match 93.2%; Score 69; DB 11; Length 454;  
 Best Local Similarity 93.3%; Pred. No. 0.00052;  
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 KONSSTTSKRRGKK 15

DB 407 KONSSTTSKRRGKK 421

RESULT 3  
 ID Q9CRY1 PRELIMINARY; PRT; 241 AA.  
 AC Q9CRY1;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE 13 days embryo head cDNA, RIKEN full-length enriched library.  
 DE clone:3100002D20, full insert sequence (fragment).  
 GN RBM6  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=EMBRYONIC HEAD;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Guelincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakomoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK013915; BAB29054.1; -  
 DR MGD: MGI:1338037; Rbm6.  
 FT NON TER 241 241  
 FT NON TER 241 241  
 SQ SEQUENCE 241 AA; 27350 MW; 115150D1E96CE71 CRC64;

Query Match 64.9%; Score 48; DB 11; Length 241;  
 Best Local Similarity 66.7%; Pred. No. 1.4; 4; Indels 0; Gaps 0;  
 Matches 10; Conservative 1; Mismatches 1;  
 QY 1 KONSSTTSKRRGKK 15  
 DB 144 KNSKTSKRRGKK 158

RESULT 4  
 ID O88720 PRELIMINARY; PRT; 1117 AA.  
 AC O88720;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE RNA binding protein.  
 GN RBM6 OR DEF-3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99280069; PubMed=10353602;  
 RA Drabkin H.A., West J., Hofflander W., Heng Y.M., Erickson P., Calvo R.,

RA Dalmu J., Gemmill R.M., Sablitzky F.,  
 RT "DEF-3, an RNA binding protein from the 3p21.3 homozygous deletion  
 region in SCLC.";  
 RT Oncogene 18:2589-2597 (1999).  
 RL (2)  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99390180; PubMed=10460589;  
 RA Hotfeller M., Sablitzky F.,  
 RT "Def-2, -3, -4 and -8, novel mouse genes differentially expressed in  
 the haemopoietic system.";  
 RT Br. J. Haematol. 106:335-344 (1999).  
 RL EMBL; AJ006486; CAA07065.2;  
 DR MGD; MGI:138037; Rbm6.  
 DR InterPro; IPR000467; G\_patch.  
 DR InterPro; IPR000504; RNA\_rec\_mot.  
 DR InterPro; IPR003955; RRM\_2.  
 DR InterPro; IPR000822; Znf\_C2H2.  
 DR Pfam; PF01585; G\_patch; 1.  
 DR Pfam; PF00076; rrm; 1.  
 DR SMART; SM00443; G\_patch; 1.  
 DR SMART; SM00360; RRM; 1.  
 DR SMART; SM00362; RRM\_2; 1.  
 DR SMART; SM00355; Znf\_C2H2; 1.  
 DR PROSITE; PS50102; RRM; 2.  
 DR PROSITE; PS50157; ZINC\_FINGER\_C2H2\_2; 2.  
 DR DNA-binding; zinc-finger.  
 KW  
 SQ SEQUENCE 1117 AA; 128092 MW; F4485367E61AB107 CRC64;

Query Match 64.9%; Score 48; DB 11; Length 1117;  
 Best Local Similarity 66.7%; Pred. No. 6.2;  
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 KONSSTTSKRGRK 15  
 DB 837 KSNKSTTSKRGRK 851

RESULT 5  
 OQR388 PRELIMINARY; PRT; 1122 AA.  
 AC OQR388;  
 DT 01-JUN-2002 (TRENBLREL. 21, Created)  
 DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)  
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)  
 DE Similar to RNA binding motif protein 6.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RA Strauberg R.,  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC026129; AAH26129.1;  
 SQ SEQUENCE 1122 AA; 128836 MW; 8FE253C8E685CD5C CRC64;

Query Match 64.9%; Score 48; DB 11; Length 1122;  
 Best Local Similarity 66.7%; Pred. No. 6.2;  
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 KONSSTTSKRGRK 15  
 DB 837 KSNKSTTSKRGRK 851

RESULT 6  
 OQXID7 PRELIMINARY; PRT; 530 AA.  
 ID OQXID7;  
 AC OQXID7;  
 DT 01-NOV-1999 (TRENBLREL. 12, Created)  
 DT 01-NOV-1999 (TRENBLREL. 12, Last sequence update)  
 DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)  
 DE Hypothetical 59.7 kDa protein F23M19.1 (late embryogenesis abundant

DE protein, putative).  
 GN F23M19.1 OR F7P12.5.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 NCBI\_TaxID=3702;  
 [1]  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV, COLUMBIA;  
 RA Vysotskaya V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,  
 RA Lee J., Liu A., Li J., Kremetska I., Luros J., Gonzalez A.,  
 RA Altafi H., Araujo R., Chao Q., Conn L., Conway A.B., Dunn P.,  
 RA Hansen N., Huizer L., Kim C., Palm C., Rowley D., Shinn P., Walker M.,  
 RA Davis R.W., Becker J.R., Federpsiel N.A., Theologis A.,  
 RT "Arabidopsis thaliana chromosome 1 BAC F23M19 sequence.";  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV, COLUMBIA;  
 RA Theologis A.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

[3]  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV, COLUMBIA;  
 RX MEDLINE=21016719; PubMed=11130712;  
 RA Theologis A., Becker J.R., Palm C.J., Federpsiel N.A., Kaul S.,  
 RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
 RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujita C.Y.,  
 RA Gall J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizer L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremetska I., Kutz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,  
 RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Utterback T., Van Aken S., Vayberg M., Vysotskaya V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 thaliana.";  
 RL Nature 408:816-820 (2000).  
 DR EMBL; AC007454; AAD39602.1;  
 DR EMBL; AC023913; AAG51902.1;  
 DR InterPro; IPR000379; Ser\_estr\_site.  
 KW Hypothetical protein.

SQ SEQUENCE 530 AA; 59670 MW; EA07EC8C787A6589 CRC64;

Query Match 62.2%; Score 46; DB 10; Length 530;  
 Best Local Similarity 53.3%; Pred. No. 6.7;  
 Matches 8; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

OY 1 KONSSTTSKRGRK 15  
 DB 460 EDDYNTTKTKGRK 474

RESULT 7  
 OQBW51 PRELIMINARY; PRT; 857 AA.  
 ID OQBW51;  
 AC OQBW51; Q9NTW0;  
 DT 01-JUN-2001 (TRENBLREL. 17, Created)  
 DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)  
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)  
 DE TTK protein kinase (D0357D13.3) (PYT).  
 GN TTK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NCBI\_TaxID=9606;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-EYE;  
 RA Strausberg R.;  
 RN Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE OF 17-857 FROM N.A.  
 RA TROMANS A.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 CC - SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; BC000633; AA00633.1; -;  
 DR EMBL; AL133475; CAB87580.1; -;  
 DR HSSP; Q00534; 1B18.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00220; S\_TKC; 1.  
 DR SMART; SM00219; TYRK; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE DOM; 1.  
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 857 AA; 97072 MW; 51F40A3CD1677AC5 CRC64;  
 Query Match 60.8%; Score 45; DB 4; Length 857;  
 Best Local Similarity 69.2%; Pred. No. 16;  
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 NSSKTTSKRRGK 15  
 DB 845 SSSKTFEKKRGK 857  
 RESULT 8  
 ID 024338 PRELIMINARY; PRT; 425 AA.  
 AC 024338;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE BSC protein (SD03549P).  
 GN BSC OR CG14941.  
 OS Drosophila melanogaster (Fruit Fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NC NCBI\_Taxid=7227;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champs M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Ball J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,  
 RA Burks D.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck U.,  
 RA Houston K.A., Howland K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris D.M., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RA The genome sequence of Drosophila melanogaster.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96016202; PubMed=7556071;  
 RA Gutjahr T., Frei E., Spicer C., Baumgartner S., White R.A., Noil M.,  
 RT "the Polycmb-group gene, extra sex combs, encodes a nuclear member of  
 the WD-40 repeat family."  
 RL EMBO J. 14:4296-4306(1995).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RA Celinker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazey R.G.,  
 RA Butenhof C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,  
 RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A., Lee B.,  
 RA Houston K.A., Hummel S.R., Karra K., Kearney L., Kim E., Kim B.,  
 RA Lewis S., Li P., Lomtan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,  
 RA Nixon K., Pacle J.M., Park S., Pfeiffer B., Poon L., Punch E.,  
 RA Sequeira A., Sethi H., Snir E., Svirskas R.R., Twomey B., Wan K.H.,  
 RA Weinburg T., Zhang R., Zieran L.L., Rubin G.M.,  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 [4]  
 RP SEQUENCE FROM N.A.  
 RA Stapleton M., Brockstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
 RA Gonzalez M., Garin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nuno C., Lewis S.E., Rubin G.M., Celinker S.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celinker S.,  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 CC - SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL; AB001634; AAP5124.1; -;  
 DR EMBL; L41867; AAA86427.1; -;  
 DR EMBL; AC006240; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; AY069796; AAL39941.1; -;  
 DR TRANSFAC; T02156; -;  
 DR FLYBase; FBgn000058; esc.  
 DR InterPro; IPR000408; Reg\_chr\_condens.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 7.  
 DR SMART; SM00320; WD40; 3.  
 DR PROSITE; PS00626; RCT\_2; UNKNOWN\_1.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; UNKNOWN\_1.  
 DR PROSITE; PSS0082; WD\_REPEATS\_2; 2.  
 DR PROSITE; PSS0294; WD\_REPEATS\_REGION; 1.  
 KM Repeat, WD repeat.  
 SQ SEQUENCE 425 AA; 47987 MW; 511G30SE5DE86727 CRC64;  
 Query Match 59.5%; Score 44; DB 5; Length 425;  
 Best Local Similarity 60.0%; Pred. No. 12;  
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 KNSKTTSKRRGK 15  
 DB 36 KSPSSTSKRRGR 50  
 RESULT 9

09SGM6 PRELIMINARY; PRT; 124 AA.  
 AC 09SGM6;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE PIN19.6.  
 OS Arabidopsis thaliana (mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eucosida II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,  
 RA Shum P., Tambunga G., Altafi H., Bai Q., Chin C., Chio J., Choi E.,  
 RA Conn L., Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B.,  
 RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharbeky N.,  
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Federspiel N.A.,  
 RA Theologis A., Becker J.R.,  
 RT "Genomic sequence for Arabidopsis thaliana BAC FIN19 from chromosome  
 1."  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; ACC09519; AF19669.1; -  
 SQ SEQUENCE 124 AA; 14188 MW; 08DAB2EC13FE3D2A CRC64;

Query Match 58.1%; Score 43; DB 10; Length 124;  
 Best Local Similarity 50.0%; Pred. No. 5.6;  
 Matches 7; Conservative 6; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 QNSKTSKRKK 15  
 DB 97 ENATKTTTCKKKK 110

RESULT 10  
 ID 09PUX7 PRELIMINARY; PRT; 270 AA.  
 AC 09PUX7;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE Mixed lineage leukemia-like protein (Fragment).  
 GN ML.  
 OS Lampyris sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 AC Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 NCBI\_TaxID=94303;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=9398697; PubMed=10468597;  
 RA Venkatesh B., Ning Y., Brenner S.,  
 RT "Late changes in spliceosomal introns define clades in vertebrate  
 evolution."  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:10267-10271(1999).  
 DR EMBL; AF137231; AAD53454.1; -  
 FT NON-TER 1  
 FT NON-TER 1  
 SQ SEQUENCE 270 AA; 28613 MW; F7A86860A2C67F87 CRC64;

Query Match 58.1%; Score 43; DB 13; Length 270;  
 Best Local Similarity 61.5%; Pred. No. 12;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 NSSKTSKRKK 15  
 DB 177 NSHKTGKRGR 189

RESULT 11  
 08RYU7

ID 08RYU7 PRELIMINARY; PRT; 556 AA.  
 AC 08RYU7;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE B1033B05.8 protein.  
 GN B1033B05.8.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 NCBI\_TaxID=33947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV; NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC  
 RT clone:B1033B05."  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF004223; BAB90658.1; -  
 SQ SEQUENCE 556 AA; 61490 MW; BAC85333021360E5 CRC64;

Query Match 58.1%; Score 43; DB 10; Length 556;  
 Best Local Similarity 50.0%; Pred. No. 24;  
 Matches 7; Conservative 6; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 KQNSKTSKRKK 14  
 DB 443 KQDAKTTNNKKGK 456

RESULT 12  
 ID 009977 PRELIMINARY; PRT; 683 AA.  
 AC 009977;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Hypothetical 76.4 kDa protein.  
 GN C17G10.6.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodermata; Caenorhabditis.  
 NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Johnson D.;  
 RT "The sequence of C. elegans cosmid C17G10."  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Waterston R.;  
 RT "Direct Submission."  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U28739; AAB93453.1; -  
 DR InterPro; IPR003582; ShKT.  
 DR SMART; SM00254; ShKT; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 683 AA; 76395 MW; F0B7E15001672E8A CRC64;

Query Match 58.1%; Score 43; DB 5; Length 683;  
 Best Local Similarity 53.3%; Pred. No. 29;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 KONSSTTSKRGK 15  
 DB 557 KKKKKKKKKKKKKKK 571

## RESULT 13

Q9VUP0 PRELIMINARY; PRT; 271 AA.  
 AC Q9VUP0; (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE Hypothetical 30.7 KDa protein.  
 GN YML041C.  
 OS Candida glabrata (Yeast) (Torulopsis glabrata).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxID=5478;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=CCN-1;  
 RA Chang S.J., Clancy C.J., Checkley M.A., Lewin A.S., Nguyen M.H.;  
 RT "A Candida glabrata gene homologous to the gene encoding a  
 RT hypothetical protein in the CAT2-AMD1 intergenic region of  
 RT Saccharomyces cerevisiae."  
 RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF177905; AAD5537.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 271 AA; 30728 MW; 0AE182D3679A0C9A CRC64;

Query Match 56.8%; Score 42; DB 3; Length 271;  
 Best Local Similarity 53.3%; Pred. No. 18;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 KONSSTTSKRGK 15  
 DB 39 KKKKKKKKKKKKKKK 53

RESULT 14  
 Q9XUG9 PRELIMINARY; PRT; 374 AA.  
 AC Q9XUG9;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE C43F9.6 protein.  
 GN C43F9.6.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 CC Rhabditidae; Peloderae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Mortimore B.J.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RA "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology."  
 RL Science 282:2012-2018(1998).  
 DR EMBL; Z82262; CAB05149.1; -.  
 DR InterPro: IPR000402; Na/K-ATPase\_beta.  
 DR Pfam; PF00287; Na\_K-ATPase; 1.  
 SQ SEQUENCE 374 AA; 41940 MW; D751B610D8BE72 CRC64;

Query Match 56.8%; Score 42; DB 5; Length 374;  
 Best Local Similarity 69.2%; Pred. No. 24;  
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KONSSTTSKRGK 13  
 DB 20 KONSSTTSKRGK 32

## RESULT 15

Q48545 PRELIMINARY; PRT; 1946 AA.  
 AC Q48545;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Protease precursor.  
 GN PRB.  
 OS Lactobacillus delbrueckii.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 CC Lactobacillaceae; Lactobacillus.  
 OX NCBI\_TaxID=1584;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=96236017; PubMed=8655480;  
 RA Gilbert C., Atlan D., Blanc B., Portailier R., Germond J.E.,  
 RA Lapiere L., Mollet B.;  
 RT "A new cell surface protease: sequencing and analysis of the prb  
 RT gene from Lactobacillus delbrueckii subsp. Dulgariusus."  
 RL J. Bacteriol. 178:3059-3065(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Germond J.E.;  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; L48487; AAC41529.1; -.  
 DR HSSP; P00782; 1SUB.  
 DR InterPro: IPR003137; PA.  
 DR InterPro: IPR000209; Peptidase\_58.  
 DR Pfam; PF02225; PA; 1.  
 DR Pfam; PF00082; Peptidase\_S8; 3.  
 DR PRINTS; PR00723; SUBTILISIN.  
 DR PROSITE; PS00136; SUBTILASIN; UNKNOWN\_1.  
 DR PROSITE; PS00138; SUBTILASIN\_SER; 1.  
 KW signal.  
 FT SIGNAL. 1 34 POTENTIAL.  
 FT CHAIN 193 1946 PROTEINASE.  
 FT SEQUENCE 1946 AA; 212315 MW; 21EF17D02E79C6A0 CRC64;

Query Match 56.8%; Score 42; DB 2; Length 1946;  
 Best Local Similarity 60.0%; Pred. No. 12+02;  
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 KONSSTTSKRGK 15  
 DB 1789 KKKKKKKKKKKKKKK 1803

Search completed: May 1, 2003, 20:37:51  
 Job time : 4.67364 secs

GenCore version 5.1.4\_p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:12:12 ; Search time 2.88703 Seconds  
(without alignments)  
692.324 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_407\_421  
Perfect score: 74  
Sequence: 1 KONSCKTSKRGKK 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	100.0	124	22	AA65552
2	74	100.0	455	21	AA18657
3	74	100.0	455	21	AA18657
4	74	100.0	455	21	AA18657
5	74	100.0	455	21	AA18657
6	74	100.0	455	21	AA18657
7	74	100.0	455	21	AA18657
8	74	100.0	455	21	AA18657
9	74	100.0	455	21	AA18657
10	74	100.0	455	21	AA18657

11	42	56.8	179	22	ABG17137	Novel human diagno
12	42	56.8	1946	19	AA47273	Lactobacillus bulg
13	41	55.4	18	21	AA165802	Kinase TRK mutant
14	41	55.4	425	22	AA29308	Human PRO polypept
15	41	55.4	734	22	AA478679	Human protein SEQ
16	41	55.4	1113	22	ABG03554	Novel human diagno
17	41	55.4	1998	22	ABG19486	Novel human diagno
18	40	54.1	78	23	ABP00883	Novel human diagno
19	40	54.1	90	22	AA488109	Human OREF protein
20	40	54.1	104	20	AA176540	Human ovary tumor
21	40	54.1	210	22	AA176540	Human ovary tumor
22	40	54.1	320	22	ABG21345	Human colon cancer
23	40	54.1	360	21	AA181178	Human prostate can
24	39	52.7	106	21	AA57182	Human prostate can
25	39	52.7	113	22	ABG25930	Human prostate can
26	39	52.7	184	22	AA175010	Human breast cancer
27	39	52.7	200	22	AA175010	Human breast cancer
28	39	52.7	242	22	AA175010	Human breast cancer
29	39	52.7	323	22	AA175010	Human breast cancer
30	39	52.7	328	22	AA175010	Human breast cancer
31	39	52.7	338	22	AA175010	Human breast cancer
32	39	52.7	395	22	AA175010	Human breast cancer
33	39	52.7	401	22	AA175010	Human breast cancer
34	39	52.7	524	22	AA175010	Human breast cancer
35	39	52.7	524	22	AA175010	Human breast cancer
36	39	52.7	627	22	AA175010	Human breast cancer
37	39	52.7	673	21	AA175010	Human breast cancer
38	39	52.7	1035	22	AA175010	Human breast cancer
39	39	52.7	1592	22	AA175010	Human breast cancer
40	39	52.7	1625	22	AA175010	Human breast cancer
41	39	52.7	1898	23	ABG07352	Human breast cancer
42	38	51.4	137	22	ABG02379	Human breast cancer
43	38	51.4	148	21	AA175010	Human breast cancer
44	38	51.4	148	21	AA175010	Human breast cancer
45	38	51.4	159	21	AA175010	Human breast cancer

## ALIGNMENTS

RESULT 1	AA65552	Standard; Protein; 124 AA.
ID	AA65552	Standard; Protein; 124 AA.
AC	AA65552	Standard; Protein; 124 AA.
XX	AA65552	Standard; Protein; 124 AA.
DT	27-MAR-2001	(first entry)
DE	C-terminal specific to novel human protein kinase MLK4B.	
KW	Human; protein kinase; antiarthritis; antileukemic; immunosuppressive;	
KW	cardiac; renal; antiinflammatory; antidiabetic; osteoporotic;	
KW	dermatological; antidiabetic; antifertility; gene therapy; vaccine;	
KW	immune disorder; cardiovascular disease; neurodegenerative disease;	
KW	cancer; autoimmune disorder; stroke; inflammatory bowel disease;	
KW	inflammatory pelvic disease; multiple sclerosis; psoriasis.	
OS	Homo sapiens.	
PN	WO200073469-A2.	
PD	07-DEC-2000.	
PF	26-MAY-2000; 2000WO-US14842.	
PR	28-MAY-1999; 99US-0136503.	
XX	(SUGEN) SUGEN INC.	
XX	Plozman GD, Martinez R, Whyte D, Sudersanam S;	
XX	WPI, 2001-032161/04.	
XX		





CC inflammatory disorder such as Addison's disease, acquired  
 CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,  
 CC rheumatoid arthritis, microbial infection and trauma.  
 XX  
 SQ Sequence 455 AA;

Query Match 100.0%; Score 74; DB 21; Length 455;  
 Best Local Similarity 100.0%; Pred. No. 0.00015;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGKK 15  
 DB 407 KONSSTTSKRRGKK 421

## RESULT 3

AAV83278  
 ID AAV83278 standard; Protein; 455 AA.  
 AAV83278;

16-AUG-2000 (first entry)  
 Human survival regulating kinase (SRK).

Survival regulating kinase; SRK; Raf; MBP; BAD; MAPKK; MAPKK;  
 HAX-1; protein kinase; autophosphorylation; cell growth; regulation;  
 apoptosis; cell survival; nuclear targeting; tumour; human;  
 autoimmune disease.

Homo sapiens.

MO200022142-A2.

20-APR-2000.

20-SEP-1999; 99WO-US22008.

13-OCT-1998; 98US-0104088.

(ONVX-) ONVX PHARM INC.

Rugieri R, Callow M, Diaz P;

WPI; 2000-317994/27.

N-PSDB; AA293783.

Novel human survival regulating kinase polypeptide for screening agents  
 which modulate biological pathways associated with SRK useful in  
 treating autoimmune diseases, tumors and apoptosis-related disorders  
 Claim 4; Figure 2; 62pp; English.

Survival regulating kinases (SRK) are a class of proteins involved in  
 cell signal transduction pathways such as mitogen-activated protein  
 kinase pathways. A protein kinase activity means that the SRK can  
 catalyze a reaction in which a phosphate group is transferred from a  
 phosphate donor to a phosphate acceptor amino acid residue.  
 CC preferably the hydroxyl side chain of a serine or threonine.  
 CC Substrates for SRK include SRK, MBP and BAD and SRK's protein kinase  
 activity is similar to that of a MAPKK such as Raf. has a range of  
 CC other activities including a cell growth-regulatory activity, a cell  
 CC survival promoting activity, a HAX-1 binding activity, an apoptosis  
 CC suppressing activity a MAPKK activation or stimulatory activity, a  
 CC nuclear targeting activity and a SRK-specific immunogenic activity.  
 CC SRK is useful for identifying agents which modulate cellular  
 CC transformations mediated by Ras and SRK and agents that modulate the  
 CC apoptosis suppression activity of SRK. This information may be useful  
 CC in the treatment of autoimmune diseases, tumors and apoptosis  
 CC related disorders.

Sequence 455 AA;

Query Match 100.0%; Score 74; DB 21; Length 455;  
 Best Local Similarity 100.0%; Pred. No. 0.00015;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGKK 15  
 DB 407 KONSSTTSKRRGKK 421.

## RESULT 4

AAV84321  
 ID AAV84321 standard; Protein; 455 AA.

AAV84321;

12-JUL-2000 (first entry)

A human cardiovascular system associated protein kinase-2.

Human; cardiovascular system associated protein kinase-2; CSAPK-2;  
 signalling pathway; cell growth; cell differentiation; gene mapping;  
 tissue typing; forensic identification; cardiovascular disease;  
 congestive heart failure; transgenic animal.

Homo sapiens.

MO200014212-A1.

16-MAR-2000.

09-SEP-1999; 99WO-US20631.

09-SEP-1998; 98US-0099657.

29-SEP-1998; 98US-0163115.

(MILL-) MILLENNIUM PHARM INC.

WPI; 2000-271053/23.

N-PSDB; AA299726; AA299727.

New nucleic acid encoding cardiovascular system associated protein  
 kinase, used e.g. for diagnosis, treatment and prevention of  
 cardiovascular disease

Claim 2; Fig 2; 163pp; English.

The present sequence represents a human cardiovascular system associated  
 CC protein kinase-2 (CSAPK-2). CSAPK polypeptides are involved in signalling  
 CC pathways associated with cell growth and differentiation. The CSAPK  
 CC polypeptides and polynucleotides are used to screen for agents that  
 CC specifically modulate CSAPK, which are potential therapeutic agents.  
 CC They are also used for diagnosis, prognosis or monitoring of  
 CC CSAPK-related diseases, gene mapping, tissue typing and forensic  
 CC identification, and for treating or preventing disorders associated  
 CC with aberrant CSAPK expression or activity, especially cardiovascular  
 CC diseases such as congestive heart failure. They can also be used in  
 CC pharmacogenomics. The CSAPK polynucleotide may also be used to generate  
 CC transgenic animals.

Sequence 455 AA;

Query Match 100.0%; Score 74; DB 21; Length 455;  
 Best Local Similarity 100.0%; Pred. No. 0.00015;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KONSSTTSKRRGKK 15  
 DB 407 KONSSTTSKRRGKK 421

## RESULT 5

Sequence 473 AA;

2 KÖNSSKITTSKRRG 14

RESULT 7  
ID ABB69961 standard; Protein; 425 AA.  
XX  
AC ABB69961;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 36675.  
XX  
KM Drosophila: developmental biology; cell signalling; insecticide;  
XX pharmacological.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PR 27-SEP-2001.  
XX  
PR 23-MAR-2001; 2001WO-US092331.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
XX  
PR 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
XX  
DR N-PSDB; ABL14064.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
PS Discloure; SEQ ID NO 36675; 21pp + Sequence listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutic and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB16176-AB160511), expressed DNA  
CC sequences (AB161840-AB161675) and the encoded proteins  
CC (ABBS7737-ABBS72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 425 AA:  
XX  
Query Match 59.5%; Score 44; DB 22; Length 425;  
Best Local Similarity 60.0%; Pred. No. 21;  
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
XX  
Oy 1 KONSXKTSKRGKK 15  
Db 36 KSPSSSTRSKRGRR 50  
XX  
RESULT 8  
ID AAG28664 standard; Protein; 51 AA.  
XX  
AC AAG28664;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 33967.  
XX  
KM Protein identification; signal transduction pathway; metabolic pathway;

KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132486.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139889.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.

```

PR 24-JUN-1999; 99US-0140695-
PR 28-JUN-1999; 99US-0140823-
PR 29-JUN-1999; 99US-0140981-
PR 30-JUN-1999; 99US-0141287-
PR 01-JUL-1999; 99US-0141842-
PR 01-JUL-1999; 99US-0142154-
PR 02-JUL-1999; 99US-0142055-
PR 06-JUL-1999; 99US-0142390-
PR 08-JUL-1999; 99US-0142803-
PR 09-JUL-1999; 99US-0142920-
PR 12-JUL-1999; 99US-0142977-
PR 13-JUL-1999; 99US-0143542-
PR 14-JUL-1999; 99US-0143624-
PR 15-JUL-1999; 99US-0144005-
PR 16-JUL-1999; 99US-0144085-
PR 16-JUL-1999; 99US-0144086-
PR 19-JUL-1999; 99US-0144325-
PR 19-JUL-1999; 99US-0144331-
PR 19-JUL-1999; 99US-0144332-
PR 19-JUL-1999; 99US-0144333-
PR 19-JUL-1999; 99US-0144334-
PR 19-JUL-1999; 99US-0144335-
PR 20-JUL-1999; 99US-0144352-
PR 20-JUL-1999; 99US-0144632-
PR 20-JUL-1999; 99US-0144884-
PR 21-JUL-1999; 99US-0144814-
PR 21-JUL-1999; 99US-0145086-
PR 21-JUL-1999; 99US-0145088-
PR 22-JUL-1999; 99US-0145085-
PR 22-JUL-1999; 99US-0145087-
PR 22-JUL-1999; 99US-0145089-
PR 22-JUL-1999; 99US-0145192-
PR 23-JUL-1999; 99US-0145145-
PR 23-JUL-1999; 99US-0145218-
PR 23-JUL-1999; 99US-0145224-
PR 26-JUL-1999; 99US-0145276-
PR 27-JUL-1999; 99US-0145918-
PR 27-JUL-1999; 99US-0145919-
PR 27-JUL-1999; 99US-0145919-
PR 28-JUL-1999; 99US-0145951-
PR 02-AUG-1999; 99US-0146386-
PR 02-AUG-1999; 99US-0146388-
PR 02-AUG-1999; 99US-0146389-
PR 03-AUG-1999; 99US-0147038-
PR 04-AUG-1999; 99US-0147204-
PR 04-AUG-1999; 99US-0147302-
PR 05-AUG-1999; 99US-0147192-
PR 05-AUG-1999; 99US-0147260-
PR 06-AUG-1999; 99US-0147303-
PR 06-AUG-1999; 99US-0147416-
PR 09-AUG-1999; 99US-0147493-
PR 09-AUG-1999; 99US-0147935-
PR 10-AUG-1999; 99US-0148171-
PR 11-AUG-1999; 99US-0148319-
PR 12-AUG-1999; 99US-0148341-
PR 13-AUG-1999; 99US-0148684-
PR 13-AUG-1999; 99US-0148684-
PR 16-AUG-1999; 99US-0149368-
PR 17-AUG-1999; 99US-0149175-
PR 18-AUG-1999; 99US-0149426-
PR 20-AUG-1999; 99US-0149722-
PR 20-AUG-1999; 99US-0149723-
PR 20-AUG-1999; 99US-0149929-
PR 23-AUG-1999; 99US-0149902-
PR 23-AUG-1999; 99US-0149930-
PR 25-AUG-1999; 99US-0150566-
PR 26-AUG-1999; 99US-0150884-
PR 27-AUG-1999; 99US-0151065-
PR 27-AUG-1999; 99US-0151066-
PR 27-AUG-1999; 99US-0151080-
PR 30-AUG-1999; 99US-0151303-
PR 31-AUG-1999; 99US-0151438-
PR 01-SEP-1999; 99US-0151930-

```

```

PR 07-SEP-1999; 99US-0152363-
PR 10-SEP-1999; 99US-0153070-
PR 13-SEP-1999; 99US-0153758-
PR 15-SEP-1999; 99US-0154018-
PR 16-SEP-1999; 99US-0154039-
PR 20-SEP-1999; 99US-0154779-
PR 22-SEP-1999; 99US-0155139-
PR 23-SEP-1999; 99US-0155486-
PR 24-SEP-1999; 99US-0155659-
PR 28-SEP-1999; 99US-0156456-
PR 29-SEP-1999; 99US-0156596-
PR 04-OCT-1999; 99US-0157117-
PR 05-OCT-1999; 99US-0157753-
PR 06-OCT-1999; 99US-0157865-
PR 07-OCT-1999; 99US-0158029-
PR 08-OCT-1999; 99US-0158232-
PR 12-OCT-1999; 99US-0158369-
PR 13-OCT-1999; 99US-0159293-
PR 13-OCT-1999; 99US-0159294-
PR 13-OCT-1999; 99US-0159295-
PR 14-OCT-1999; 99US-0159329-
PR 14-OCT-1999; 99US-0159330-
PR 14-OCT-1999; 99US-0159331-
PR 14-OCT-1999; 99US-0159637-
PR 14-OCT-1999; 99US-0159638-
PR 18-OCT-1999; 99US-0159584-
PR 21-OCT-1999; 99US-0160741-
PR 21-OCT-1999; 99US-0160767-
PR 21-OCT-1999; 99US-0160768-
PR 21-OCT-1999; 99US-0160770-
PR 21-OCT-1999; 99US-0160814-
PR 21-OCT-1999; 99US-0160815-
PR 22-OCT-1999; 99US-0160980-
PR 22-OCT-1999; 99US-0160981-
PR 22-OCT-1999; 99US-0160989-
PR 25-OCT-1999; 99US-0161404-
PR 25-OCT-1999; 99US-0161405-
PR 25-OCT-1999; 99US-0161406-
PR 26-OCT-1999; 99US-0161359-
PR 26-OCT-1999; 99US-0161360-
PR 26-OCT-1999; 99US-0161361-
PR 26-OCT-1999; 99US-0161920-
PR 28-OCT-1999; 99US-0161992-
PR 28-OCT-1999; 99US-0161993-
PR 29-OCT-1999; 99US-0162142-

```

```

Query Match 58.1%; Score 43; DB 21; Length 51;
Best Local Similarity 50.0%; Pred. No. 3.3;
Matches 7; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 2 QNSKTTSKRRGK 15
Db 24 ENATKTTTXXXXX 37

```

```

RESULT 9
ID AAG41571 standard; Protein; 124 AA.
AAAG41571

```

```

XX AAG41571;
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 51740.

```

```

XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX

```

PD 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-0301439.  
PF 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0122548.  
PR 23-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129844.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0131089.  
PR 30-APR-1999; 99US-0131444.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136332.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140685.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145376.  
PR 27-JUL-1999; 99US-0145313.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 17-AUG-1999; 99US-0149368.  
PR 18-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.

PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 25-OCT-1999; 99US-0161407.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 58.1%; Score 43; DB 21; Length 124;  
 Best Local Similarity 50.0%; Pred. No. 8.6;  
 Matches 7; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 2 QNSKTTSKRRGK 15  
 DB 97 ENATKTTTKKKKK 110

RESULT 10  
 ABG14321  
 ABG14321 standard; Protein; 179 AA.

XX ABG14321;  
 AC 18-FEB-2002 (first entry)  
 DT XX  
 DE Novel human diagnostic protein #14312.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200175067-A2.  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US08631.  
 XX  
 PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX  
 PA (HYSE-) HYSEQ INC.

XX  
 PI Drmanac RT, Liu C, Tang YT;  
 DR WPI; 2001-639362/73.  
 DR N-PSDB; AAS78508.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostic, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity  
 XX  
 PS Claim 20; SEQ ID No 44680; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences (I) is useful as hybridisation probes,  
 CC polynucleotide chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pat\_sequences.

SQ Sequence 179 AA;

Query Match 56.8%; Score 42; DB 22; Length 179;  
 Best Local Similarity 57.1%; Pred. No. 19;  
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 QNSKTTSKRRGK 15  
 DB 11 QNSSETIKKNGKK 24

RESULT 11  
 ABG17137  
 ID ABG17137 standard; Protein; 179 AA.

XX ABG17137;  
 AC 18-FEB-2002 (first entry)  
 DT XX  
 DE Novel human diagnostic protein #17128.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200175067-A2.  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US08631.  
 XX  
 PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.  
 DR N-PSDB; AAS81324.

PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -

PS Claim 20; SEQ ID No 47496; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations in  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

CC Sequence 179 AA;

Query Match 56.8%; Score 42; DB 22; Length 179;  
 Best Local Similarity 57.1%; Pred. No. 19;  
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 QNSKTTSKRGKK 15  
 DB 11 QNSSETIKNGKK 24

RESULT 12  
 AAW47273  
 ID AAW47273 standard; Protein; 1946 AA.

XX AAW47273;

DT 22-MAY-1998 (first entry)

DE Lactobacillus bulgaricus protease P.

KW Protease P; transformed microorganism; recombinant production;  
 KW Prp; yoghurt; acidulated milk; cheese.

OS Lactobacillus bulgaricus.

PN EP810289-A1.

PD 03-DEC-1997.

PF 29-MAY-1996; 96EP-0201495.

PR 29-MAY-1996; 96EP-0201495.

PA (NEST ) SOC PROD NESTLE SA.

PI Germond JE, Lapierre L, Mollet B;

DR WPI; 1998-011060/02.  
 DR N-PSDB; AAV15586.

XX Recombinant Lactobacillus bulgaricus Prp protease - and DNA for  
 PT transforming microorganisms for making fermented dairy products  
 PT XX  
 PS Claim 1; Pages 9-17; 28pp; English.

CC The present sequence is Lactobacillus bulgaricus protease P  
 CC (Prp). A microorganism transformed with the Prp DNA can be used  
 CC for the recombinant production of the enzyme, or in the manufacture  
 CC of fermented dairy products, e.g. yoghurt, acidulated milk and  
 CC cheese, especially where: (a) the microorganism expresses a  
 CC recombinant Prp derivative that is temperature or pH sensitive,  
 CC preferably having no more than 80% of wild type protease activity  
 CC under storage conditions and at least 90% of wild type protease  
 CC activity under fermentation conditions; or (b) the microorganism  
 CC contains Prp DNA under the control of a temperature or pH  
 CC sensitive promoter.

CC Sequence 1946 AA;

Query Match 56.8%; Score 42; DB 19; Length 1946;  
 Best Local Similarity 60.0%; Pred. No. 2,4e+02;  
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 KNSKTTSKRGKK 15  
 DB 1789 KKTDSKTTSSAKK 1803

RESULT 13  
 AAY65802  
 ID AAY65802 standard; Peptide; 18 AA.

XX AAY65802;

DT 10-FEB-2000 (first entry)

DE Kinase TTK mutant peptide 11.

KW Human; frameshift mutant; T cell response; tumour; treatment; cancer;  
 KW mutein.

OS Homo sapiens.  
 OS Synthetic.

PN WO9958552-A2.

PD 18-NOV-1999.

PF 03-MAY-1999; 99WO-NO00143.

PR 08-MAY-1998; 98NO-0002097.

PA (NHHD ) NORSK HYDRO AS.

PI Gaudernack G, Eriksen JA, Moller M, Gjertsen MK, Sæterdal I;  
 DR WPI; 2000-039064/03.

PT New peptides derived from genes with frameshift mutations, used to  
 PT develop products for the treatment and prophylaxis of cancers -  
 PS Claim 13; Page 27; 166pp; English.

CC Peptides AAY6584-Y66142 are fragments of mutant proteins arising from a  
 CC frameshift mutation in a gene from a cancer cell. The peptides are  
 CC characterised in that they:  
 CC (i) are at least 8 amino acids long and a fragment of a mutant protein  
 CC arising from a frameshift mutation in a gene of a cancer cell;  
 CC (ii) consist of at least one amino acid of the mutant part of a protein  
 CC sequence encoded by the gene;  
 CC (iii) comprise 0-10 amino acid from the carboxyl terminus of the normal  
 CC part of the protein sequence preceding the amino terminus of the mutant





```

XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
DN (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhtman T, Goodrich R;
XX
DR WP1; 2001-476283/51.
XX
DR N-PSDB; AAK51812.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy.
XX
PS Claim 20; Page 3585-3586; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAW78323-AAW80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAW80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
Sequence 734 AA;
Query Match 55.4%; Score 41; DB 22; Length 734;
Best Local Similarity 53.3%; Pred. NO. 1.3e+02;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 KONSSTTSKRRGKK 15
|:|||||:|:|
DB 552 KEKASKTTVRRRKR 566

```

Search completed: May 1, 2003, 20:35:38  
 Job time : 4.86703 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using BW model

Run on: May 1, 2003, 20:38:08 ; Search time 0.962343 Seconds  
(without alignments)  
1344.947 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_407\_421

Perfect score: 74

Sequence: 1 KONSSTKTSKRGRKK 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328255 seqs, 8628685 residues

Optimal number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PTCT\_NEW\_PUB pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PTCTUS\_PUBCOMB pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	100.0	455	10 US-09-757-982-5	Sequence 5, Appl1
2	45	60.8	841	9 US-10-081-119-14	Sequence 14, Appl1
3	42	56.8	312	10 US-09-801-368-260	Sequence 260, App
4	41	55.4	425	9 US-10-174-590-570	Sequence 570, App
5	41	55.4	425	9 US-10-176-758-570	Sequence 570, App
6	41	55.4	425	9 US-10-175-737-570	Sequence 570, App
7	41	55.4	425	9 US-10-173-706-570	Sequence 570, App
8	41	55.4	425	9 US-10-175-738-570	Sequence 570, App
9	41	55.4	425	9 US-10-175-752-570	Sequence 570, App
10	41	55.4	425	9 US-10-176-482-570	Sequence 570, App
11	41	55.4	425	9 US-10-176-757-570	Sequence 570, App
12	41	55.4	425	9 US-10-176-913-570	Sequence 570, App
13	41	55.4	425	9 US-10-180-552-570	Sequence 570, App
14	41	55.4	425	9 US-10-180-557-570	Sequence 570, App
15	41	55.4	425	9 US-10-173-700-570	Sequence 570, App
16	41	55.4	425	9 US-10-174-572-570	Sequence 570, App
17	41	55.4	425	9 US-10-174-579-570	Sequence 570, App
18	41	55.4	425	9 US-10-174-582-570	Sequence 570, App
19	41	55.4	425	9 US-10-174-588-570	Sequence 570, App

20	41	55.4	425	9	US-10-175-739-570	Sequence 570, App
21	41	55.4	425	9	US-10-175-740-570	Sequence 570, App
22	41	55.4	425	9	US-10-175-743-570	Sequence 570, App
23	41	55.4	425	9	US-10-176-488-570	Sequence 570, App
24	41	55.4	425	9	US-10-176-492-570	Sequence 570, App
25	41	55.4	425	9	US-10-176-747-570	Sequence 570, App
26	41	55.4	425	9	US-10-176-750-570	Sequence 570, App
27	41	55.4	425	9	US-10-176-985-570	Sequence 570, App
28	41	55.4	425	9	US-10-176-987-570	Sequence 570, App
29	41	55.4	425	9	US-10-176-991-570	Sequence 570, App
30	41	55.4	425	9	US-10-176-992-570	Sequence 570, App
31	41	55.4	425	9	US-10-174-576-570	Sequence 570, App
32	41	55.4	425	9	US-10-184-658-570	Sequence 570, App
33	41	55.4	425	9	US-10-173-695-570	Sequence 570, App
34	41	55.4	425	9	US-10-173-697-570	Sequence 570, App
35	41	55.4	425	9	US-10-173-705-570	Sequence 570, App
36	41	55.4	425	9	US-10-174-576-570	Sequence 570, App
37	41	55.4	425	9	US-10-174-585-570	Sequence 570, App
38	41	55.4	425	9	US-10-174-586-570	Sequence 570, App
39	41	55.4	425	9	US-10-175-747-570	Sequence 570, App
40	41	55.4	425	9	US-10-176-481-570	Sequence 570, App
41	41	55.4	425	9	US-10-176-485-570	Sequence 570, App
42	41	55.4	425	9	US-10-176-487-570	Sequence 570, App
43	41	55.4	425	9	US-10-176-493-570	Sequence 570, App
44	41	55.4	425	9	US-10-176-756-570	Sequence 570, App
45	41	55.4	425	9	US-10-176-911-570	Sequence 570, App

## ALIGNMENTS

RESULT 1  
US-09-757-982-5  
Sequence 5, Application US/09757982  
Patent No. US20020094559A1  
GENERAL INFORMATION:  
APPLICANT: Acton, Susan  
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
FILE REFERENCE: MNI-050  
CURRENT APPLICATION NUMBER: US/09/757,982  
CURRENT FILING DATE: 2001-01-10  
PRIOR APPLICATION NUMBER: 09/163,115  
PRIOR FILING DATE: 1998-09-29  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 455  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-757-982-5

Query Match 100.0%; Score 74; DB 10; Length 455;  
Best Local Similarity 100.0%; Pred. No. 0.00097;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KONSSTKTSKRGRKK 15  
Db 407 KONSSTKTSKRGRKK 421

RESULT 2  
US-10-081-119-14  
Sequence 14, Application US/10081119  
Publication No. US20030045491A1  
GENERAL INFORMATION:  
APPLICANT: Reinhard, Christoph  
APPLICANT: Jefferson, Anne B.  
TITLE OF INVENTION: TTK in Diagnosis and as a Therapeutic  
FILE REFERENCE: 16932.002  
CURRENT APPLICATION NUMBER: US/10/081,119  
CURRENT FILING DATE: 2002-02-21

PRIOR APPLICATION NUMBER: 60/289,813  
PRIOR FILING DATE: 2001-02-21  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14  
LENGTH: 841  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-081-119-14

Query Match 60.8%; Score 45; DB 9; Length 841;  
Best Local Similarity 69.2%; Pred. No. 42;  
Matches 9; Conservative 2; Mismatches 0; Gaps 0;

QY 3 NSSKTSKRRGK 15  
|:|||||:|  
Db 829 SSKTFEKKRKK 841

RESULT 3

Sequence 260, Application US/09801368  
Patent No. US20020128250A1

GENERAL INFORMATION:  
APPLICANT: Busby, Robert  
APPLICANT: Call, Brian  
APPLICANT: Hecht, Peter  
APPLICANT: Holtzman, Doug  
APPLICANT: Madden, Kevin  
APPLICANT: Maxon, Mary  
APPLICANT: Milne, Todd  
APPLICANT: No. US20020128250A1man, Thea  
APPLICANT: Royer, John  
APPLICANT: Salama, Sofie  
APPLICANT: Sherman, Amir  
APPLICANT: Silva, Jeff  
APPLICANT: Summers, Eric  
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi  
FILE REFERENCE: 109272.147  
CURRENT APPLICATION NUMBER: US/09/801,368  
CURRENT FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 09/487,558  
PRIOR FILING DATE: 2000-01-19  
PRIOR APPLICATION NUMBER: US 60/160,587  
PRIOR FILING DATE: 1999-10-20  
NUMBER OF SEQ ID NOS: 440  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 260  
LENGTH: 312  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-801-368-260

Query Match 56.8%; Score 42; DB 10; Length 312;  
Best Local Similarity 64.3%; Pred. No. 45;  
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KONSSTTSKRRGK 14  
|:|||||:|  
Db 148 KSNSSPYLNKRKK 161

RESULT 4

US-10-174-590-570  
Sequence 570, Application US/10174590  
Publication No. US20030008352A1

GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jitan  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C42  
CURRENT APPLICATION NUMBER: US/10/174,590  
CURRENT FILING DATE: 2002-06-18  
Prior application removed - See file wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 570  
LENGTH: 425  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-174-590-570

Query Match 55.4%; Score 41; DB 9; Length 425;  
Best Local Similarity 53.3%; Pred. No. 86;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KONSSTTSKRRGK 15  
|:|||||:|  
Db 243 KKKSKTTVRRRRK 257

RESULT 5

US-10-176-758-570  
Sequence 570, Application US/10176758  
Publication No. US20030008353A1

GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jitan  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C104  
CURRENT APPLICATION NUMBER: US/10/176,758  
CURRENT FILING DATE: 2002-06-21  
Prior application removed - See file wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 570  
LENGTH: 425  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-176-758-570

Query Match 55.4%; Score 41; DB 9; Length 425;  
Best Local Similarity 53.3%; Pred. No. 86;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KONSSTTSKRRGK 15  
|:|||||:|  
Db 243 KKKSKTTVRRRRK 257

RESULT 6

US-10-175-737-570  
Sequence 570, Application US/10175737  
Publication No. US20030013153A1

GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jitan  
APPLICANT: Desnoyers, Luc

```
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C50
CURRENT APPLICATION NUMBER: US/10/175,737
Pilot Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 570
LENGTH: 425
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-737-570
```

```
Query Match
Best Local Similarity 55.4%; Score 41; DB 9; Length 425;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 KONSSTTSKRGRKK 15
Db 243 KEKASKTTTVRRRR 257
```

## RESULT 7

```
US-10-173-706-570
Sequence 570, Application US/10173706
Publication No. US20030022293A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C7
CURRENT APPLICATION NUMBER: US/10/173,706
Pilot Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 570
LENGTH: 425
TYPE: PRT
ORGANISM: Homo Sapien
US-10-173-706-570
```

```
Query Match
Best Local Similarity 55.4%; Score 41; DB 9; Length 425;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 KONSSTTSKRGRKK 15
Db 243 KEKASKTTTVRRRR 257
```

```
RESULT 8
US-10-175-738-570
Sequence 570, Application US/10175738
Publication No. US20030022294A1
GENERAL INFORMATION:
```

```
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C45
CURRENT APPLICATION NUMBER: US/10/175,738
Pilot Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 570
LENGTH: 425
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-738-570
```

```
Query Match
Best Local Similarity 55.4%; Score 41; DB 9; Length 425;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 KONSSTTSKRGRKK 15
Db 243 KEKASKTTTVRRRR 257
```

## RESULT 9

```
US-10-175-752-570
Sequence 570, Application US/10175752
Publication No. US20030022295A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C60
CURRENT APPLICATION NUMBER: US/10/175,752
Pilot Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 570
LENGTH: 425
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-752-570
```

```
Query Match
Best Local Similarity 55.4%; Score 41; DB 9; Length 425;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 KONSSTTSKRGRKK 15
Db 243 KEKASKTTTVRRRR 257
```

```
RESULT 10
US-10-176-482-570
```

Sequence 570, Application US/10176482  
Publication No. US20030022296A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C70  
CURRENT FILING DATE: 2002-06-20  
CURRENT APPLICATION NUMBER: US/10/176,482  
Prior Application removed - See file wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 570  
LENGTH: 425  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-176-482-570

Query Match 55.4%; Score 41; DB 9; Length 425;  
Best Local Similarity 53.3%; Pred. No. 86;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KONSSTTSKRGGK 15  
DB 243 KKKSKTTTVRRRR 257

## RESULT 11

US-10-176-757-570  
Sequence 570, Application US/10176757  
Publication No. US20030022297A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C6  
CURRENT FILING DATE: 2002-06-20  
CURRENT APPLICATION NUMBER: US/10/176,757  
Prior Application removed - See file wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 570  
LENGTH: 425  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-176-757-570

Query Match 55.4%; Score 41; DB 9; Length 425;  
Best Local Similarity 53.3%; Pred. No. 86;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KONSSTTSKRGGK 15  
DB 243 KKKSKTTTVRRRR 257

RESULT 12  
US-10-176-913-570  
Sequence 570, Application US/10176913  
Publication No. US20030022298A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C66  
CURRENT FILING DATE: 2002-06-20  
CURRENT APPLICATION NUMBER: US/10/176,913  
Prior Application removed - See file wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 570  
LENGTH: 425  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-176-913-570

Query Match 55.4%; Score 41; DB 9; Length 425;  
Best Local Similarity 53.3%; Pred. No. 86;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KONSSTTSKRGGK 15  
DB 243 KKKSKTTTVRRRR 257

RESULT 13  
US-10-180-552-570  
Sequence 570, Application US/10180552  
Publication No. US20030022300A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian  
APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3430R1C153  
CURRENT FILING DATE: 2002-06-25  
CURRENT APPLICATION NUMBER: US/10/180,552  
Prior Application removed - See file wrapper or Palm  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 570  
LENGTH: 425  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-180-552-570

Query Match 55.4%; Score 41; DB 9; Length 425;  
Best Local Similarity 53.3%; Pred. No. 86;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KONSSTTSKRGGK 15

Db 243 KEKASKTTTVRRRRK 257

```

RESULT 14
US-10-180-557-570
; Sequence 570, Application US/10180557
; Publication No. US20030022301A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C147
; CURRENT APPLICATION NUMBER: US/10/180,557
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 570
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-557-570

```

```

Query Match
Best Local Similarity 55.4%; Score 41; DB 9; Length 425;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 KONSSTTSKRGKK 15
Db 243 KEKASKTTTVRRRRK 257

```

```

RESULT 15
US-10-173-700-570
; Sequence 570, Application US/10173700
; Publication No. US20030027262A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C14
; CURRENT APPLICATION NUMBER: US/10/173,700
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 570
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-700-570

```

```

Query Match
Best Local Similarity 55.4%; Score 41; DB 9; Length 425;
Best Local Similarity 53.3%; Score 36; DB 9; Length 425;

```

```

Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 KONSSTTSKRGKK 15
Db 243 KEKASKTTTVRRRRK 257

```

```

Search completed: May 1, 2003, 20:52:15
Job time: 1.96234 secs

```





GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using BW model

Run on: May 1, 2003, 20:33:12 ; Search time 1.00418 Seconds  
(without alignments)  
439.505 Million cell updates/sec

Title: US-09-757-982-5\_COPY\_407\_421  
Perfect score: 74  
Sequence: 1 KONSSTTSKRGRK 15

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Archived: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A COMB.pep: \*  
2: /cgn2\_6/ptodata/1/1aa/5B COMB.pep: \*  
3: /cgn2\_6/ptodata/1/1aa/6A COMB.pep: \*  
4: /cgn2\_6/ptodata/1/1aa/6B COMB.pep: \*  
5: /cgn2\_6/ptodata/1/1aa/PTUS COMB.pep: \*  
6: /cgn2\_6/ptodata/1/1aa/backfilest.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	74	100.0	455	3	US-09-221-235-5
2	74	100.0	455	3	US-09-221-928-5
3	74	100.0	455	4	US-09-221-527-5
4	74	100.0	455	4	US-09-221-527-5
5	74	100.0	455	4	US-09-221-236-5
6	74	100.0	455	4	US-09-221-245-5
7	74	100.0	455	4	US-09-221-245-5
8	74	100.0	455	4	US-09-221-528-5
9	74	100.0	455	4	US-09-221-528-5
10	74	100.0	455	4	US-09-221-528-5
11	39	52.7	853	4	US-09-078-347A-2
12	39	51.4	853	4	US-07-956-483C-14
13	37	50.0	30	2	US-08-248-839C-80
14	37	50.0	211	2	US-08-708-958-2
15	37	50.0	357	1	US-08-411-777-9
16	37	50.0	357	1	US-08-057-068-9
17	37	50.0	376	1	US-08-090-013-4
18	37	50.0	376	1	US-08-081-328-4
19	37	50.0	376	1	US-08-232-249-4
20	37	50.0	376	1	US-08-833-642A-4
21	37	50.0	376	2	US-08-389-423-4
22	37	50.0	376	4	US-09-189-028-4
23	37	50.0	823	4	US-09-651-656-19
24	37	50.0	823	4	US-09-650-855-19
25	36	48.6	324	1	US-08-323-531-4
26	36	48.6	324	1	US-08-198-094-4
27	36	48.6	324	4	US-08-107-794A-4

28	36	48.6	324	5	PCT-US93-07424-4	Sequence 4, Appli
29	36	48.6	324	5	PCT-US95-02087-4	Sequence 4, Appli
30	36	48.6	401	4	US-09-651-656-103	Sequence 103, App
31	36	48.6	401	4	US-09-650-855-103	Sequence 103, App
32	36	48.6	418	2	US-08-978-182-5	Sequence 5, Appli
33	36	48.6	418	2	US-09-205-181-5	Sequence 5, Appli
34	36	48.6	602	4	US-09-134-001C-4416	Sequence 4, Appli
35	36	48.6	670	1	US-08-363-208-2	Sequence 2, Appli
36	36	48.6	670	4	US-09-137-478-2	Sequence 8, Appli
37	36	48.6	671	2	US-08-426-125-8	Sequence 8, Appli
38	36	48.6	671	2	US-08-455-355-8	Sequence 8, Appli
39	35.5	48.0	934	3	US-08-929-329-6	Sequence 6, Appli
40	35	47.3	118	3	US-09-012-084-4	Sequence 4, Appli
41	35	47.3	304	4	US-09-134-001C-4578	Sequence 4, Appli
42	35	47.3	338	2	US-09-047-026A-2	Sequence 2, Appli
43	35	47.3	534	4	US-09-198-603C-6	Sequence 6, Appli
44	35	47.3	798	4	US-09-134-001C-4917	Sequence 4, Appli
45	35	47.3	1239	2	US-08-937-931-2	Sequence 2, Appli

## ALIGNMENTS

RESULT 1  
US-09-221-235-5  
; Sequence 5, Application US/09221235  
; Patent No. 6043040  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: NMT-050  
; CURRENT APPLICATION NUMBER: US/09/221,235  
; EARLIER FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-235-5

Query Match 100.0%; Score 74; DB 3; Length 455;  
Best Local Similarity 100.0%; Pred. No. 3.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KONSSTTSKRGRK 15  
DB 407 KONSSTTSKRGRK 421

RESULT 2  
US-09-221-928-5  
; Sequence 5, Application US/09221928  
; Patent No. 6121030  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: NMT-050  
; CURRENT APPLICATION NUMBER: US/09/221,928  
; EARLIER FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-928-5

Query Match 100.0%; Score 74; DB 3; Length 455;

Best Local Similarity 100.0%; Pred. No. 3.1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KONSSTTSKRRGKK 15  
|||||  
Db 407 KONSSTTSKRRGKK 421

RESULT 3  
US-09-221-527-5

; Sequence 5, Application US/09221527

; Patent No. 6146832

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: NMI-050

; CURRENT APPLICATION NUMBER: US/09/221,527

; CURRENT FILING DATE: 1998-12-28

; EARLIER APPLICATION NUMBER: 09/163,115

; EARLIER FILING DATE:

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5

; LENGTH: 455

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-221-527-5

Query Match

Best Local Similarity 100.0%; Score 74; DB 4; Length 455;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KONSSTTSKRRGKK 15  
|||||  
Db 407 KONSSTTSKRRGKK 421

RESULT 4  
US-09-221-236-5

; Sequence 5, Application US/09221236

; Patent No. 6146841

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: NMI-050

; CURRENT APPLICATION NUMBER: US/09/221,236

; CURRENT FILING DATE: 1998-12-28

; EARLIER APPLICATION NUMBER: 09/163,115

; EARLIER FILING DATE: 1998-09-29

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5

; LENGTH: 455

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-221-236-5

Query Match

Best Local Similarity 100.0%; Score 74; DB 4; Length 455;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KONSSTTSKRRGKK 15  
|||||  
Db 407 KONSSTTSKRRGKK 421

RESULT 5  
US-09-221-416-5

; Sequence 5, Application US/09221416

; Patent No. 6151417

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: NMI-050

; CURRENT APPLICATION NUMBER: US/09/221,416

; CURRENT FILING DATE: 1998-12-28

; EARLIER APPLICATION NUMBER: 09/163,115

; EARLIER FILING DATE: 1998-09-29

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5

; LENGTH: 455

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-221-416-5

Query Match

Best Local Similarity 100.0%; Score 74; DB 4; Length 455;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KONSSTTSKRRGKK 15  
|||||  
Db 407 KONSSTTSKRRGKK 421

RESULT 6  
US-09-221-245-5

; Sequence 5, Application US/09221245

; Patent No. 6180358

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: NMI-050

; CURRENT APPLICATION NUMBER: US/09/221,245

; CURRENT FILING DATE: 1998-12-28

; EARLIER APPLICATION NUMBER: US 09/163,115

; EARLIER FILING DATE: 1998-09-29

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5

; LENGTH: 455

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-221-245-5

Query Match

Best Local Similarity 100.0%; Score 74; DB 4; Length 455;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KONSSTTSKRRGKK 15  
|||||  
Db 407 KONSSTTSKRRGKK 421

RESULT 7  
US-09-163-115-5

; Sequence 5, Application US/09163115A

; Patent No. 6183962

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: NMI-050

; CURRENT APPLICATION NUMBER: US/09/163,115A

; CURRENT FILING DATE: 1998-09-29

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 5

; LENGTH: 455

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-163-115-5

Query Match

Best Local Similarity 100.0%; Score 74; DB 4; Length 455;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KONSSTTSKRGKK 15  
|||||  
Db 407 KONSSTTSKRGKK 421

## RESULT 8

US-09-528-5  
Sequence 5, Application US/09221528  
Patent No. 6190874  
GENERAL INFORMATION:  
APPLICANT: Acton, Susan  
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
FILE REFERENCE: MNI-050  
CURRENT APPLICATION NUMBER: US/09/221,528  
EARLIER FILING DATE: 1998-12-28  
PRIOR APPLICATION NUMBER: 09/163,115  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 5  
LENGTH: 455  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-221-528-5

Query Match 100.0%; Score 74; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 3,1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KONSSTTSKRGKK 15  
|||||  
Db 407 KONSSTTSKRGKK 421

## RESULT 9

US-09-593-553-5  
Sequence 5, Application US/09593553  
Patent No. 6200770  
GENERAL INFORMATION:  
APPLICANT: Acton, Susan  
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
FILE REFERENCE: MNI-050  
CURRENT APPLICATION NUMBER: US/09/593,553  
PRIOR FILING DATE: 2000-06-14  
PRIOR APPLICATION NUMBER: 09/163,115  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 5  
LENGTH: 455  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-593-553-5

Query Match 100.0%; Score 74; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 3,1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KONSSTTSKRGKK 15  
|||||  
Db 407 KONSSTTSKRGKK 421

## RESULT 10

US-09-221-237-5  
Sequence 5, Application US/09221237  
Patent No. 6214597  
GENERAL INFORMATION:  
APPLICANT: Acton, Susan  
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
FILE REFERENCE: MNI-050  
CURRENT APPLICATION NUMBER: US/09/221,237  
CURRENT FILING DATE: 1998-12-28

EARLIER APPLICATION NUMBER: 09/163,115  
EARLIER FILING DATE: 1998-09-29  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 5  
LENGTH: 455  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-221-237-5

Query Match 100.0%; Score 74; DB 4; Length 455;  
Best Local Similarity 100.0%; Pred. No. 3,1e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KONSSTTSKRGKK 15  
|||||  
Db 407 KONSSTTSKRGKK 421

## RESULT 11

US-09-078-347A-2  
Sequence 2, Application US/09078347A  
Patent No. 6132968  
GENERAL INFORMATION:  
APPLICANT: Le, Xiao-Chun  
APPLICANT: Weinfeld, Michael  
TITLE OF INVENTION: Method for Quantitating Low Level  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Medien & Carroll, LLP  
STREET: 220 Montgomery Street, Suite 2200  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94104

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/078,347A  
FILING DATE: 13-MAY-1998  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Peter G.  
REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: UALB-03283  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 673 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-09-078-347A-2

Query Match 52.7%; Score 39; DB 4; Length 673;  
Best Local Similarity 61.5%; Pred. No. 56;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 QNSSTTSKRGKK 14  
||:|:|:|  
Db 601 QNTAKTKRGKK 613

RESULT 12  
US-07-956-483-14

Sequence 14, Application US/07956483  
Patent No. 6261799  
GENERAL INFORMATION:  
APPLICANT: KIENY, Marie-Paule  
TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED  
TITLE OF INVENTION: gp160 VARIANT  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/956,483  
FILING DATE: 31-DEC-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 92/19742  
FILING DATE: 12-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 91 05392  
FILING DATE: 02-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Crane-Feury, Sharon E  
REGISTRATION NUMBER: 36,113  
REFERENCE/DOCKET NUMBER: 017753-005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 855 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-07-956-483-14

Query Match 51.4%; Score 38; DB 4; Length 855;  
Best Local Similarity 58.3%; Pred. No. 1.1e+02;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

3 NSSKTSKRGRK 14  
DB 140 NATNTTSSNRK 151

RESULT 13  
US-08-248-839C-80  
Sequence 80, Application US/08248839C  
Patent No. 5843702  
GENERAL INFORMATION:  
APPLICANT: McConnell, David  
APPLICANT: Devine, Kevin  
APPLICANT: O'Kane, Charles  
TITLE OF INVENTION: A Gene Expression System  
NUMBER OF SEQUENCES: 185  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 58437020 No. 5843702disk of No. 5843702th America, Inc.  
STREET: 405 Lexington Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/248,839C  
FILING DATE: 25-MAY-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Gregg, Valeria A.  
REGISTRATION NUMBER: 35,127  
REFERENCE/DOCKET NUMBER: 3614.214-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123  
TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 80:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
US-08-248-839C-80

Query Match 50.4%; Score 37; DB 2; Length 30;  
Best Local Similarity 72.7%; Pred. No. 4.3;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KNSKTSKR 11  
DB 10 KNTSKTESLR 20

RESULT 14  
US-08-708-958-2  
Sequence 2, Application US/08708958  
Patent No. 5948952  
GENERAL INFORMATION:  
APPLICANT: SANDS, Arthur T.  
APPLICANT: BRADLEY, Allan  
APPLICANT: ABUIN, Alejandro  
TITLE OF INVENTION: XERODERMA PIGMENTOSUM-DEFICIENT  
TITLE OF INVENTION: MOUSE  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS  
STREET: 2100 PENNSYLVANIA AVENUE, N.W.  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/708,958  
FILING DATE: SEP-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: KIT, Gordon  
REGISTRATION NUMBER: 30,764  
REFERENCE/DOCKET NUMBER: A-6641  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 293-7060  
TELEFAX: (202) 293-7860  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 211 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-708-958-2

Query Match 50.0%; Score 37; DB 2; Length 211;  
Best Local Similarity 58.3%; Pred. No. 35;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 SSKTTSKRGRKK 15  
||:|:|:|:|  
DB 88 SSSSSSSSKRGRKK 99

RESULT 15

US-08-411-777-9

/ Sequence 9, Application US/08411777  
/ Patent No. 5792641

/ GENERAL INFORMATION:

/ APPLICANT: Schultein, Martin

/ APPLICANT: Fredholm, Henrik

/ APPLICANT: Hjorth, Carsten

/ APPLICANT: Rasmussen, Grethe

/ APPLICANT: Nielsen, Egon

/ APPLICANT: Rosholm, Peter

/ TITLE OF INVENTION: Cellulase Variants

/ NUMBER OF SEQUENCES: 10

/ CORRESPONDENCE ADDRESS: 10

/ ADDRESS: No. 57926410 No. 5792641disk of No. 5792641th America

/ STREET: 405 Lexington Avenue

/ CITY: New York

/ STATE: NY

/ COUNTRY: USA

/ ZIP: 10174-6401

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Diskette

/ COMPUTER: IBM Compatible

/ OPERATING SYSTEM: DOS

/ SOFTWARE: FastSeq for Windows Version 2.0

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/411,777

/ FILING DATE: 05-MAY-1995

/ CLASSIFICATION: 435

/ PRIORITY APPLICATION DATA:

/ APPLICATION NUMBER:

/ FILING DATE:

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Gregg, Valera A

/ REGISTRATION NUMBER: 35,127

/ REFERENCE/DOCKET NUMBER: 3913.504-US

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: 212-867-0123

/ TELEFAX: 212-868-9655

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9

/ INFORMATION FOR SEQ ID NO: 9:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 357 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: linear

/ US-08-411-777-9



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:51:28 ; Search time 47 Seconds

(without alignments)  
930.663 Million cell updates/sec

Title: US-09-757-982-5

Sequence: 1 MSSLGASFWQIKFDDIQFFE.....GDDDDDDDEEDNDNDNDE 455

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	2.4	848	2	B87950
2	11	2.4	855	2	T20082
3	10	2.2	729	2	A28651
4	10	2.2	1089	2	S48244
5	9	2.0	297	2	T30613
6	9	2.0	375	2	A39622
7	9	2.0	408	2	T08069
8	9	2.0	502	2	JC4337
9	9	2.0	502	2	I48241
10	9	2.0	502	2	S64951
11	9	2.0	856	2	T52415
12	9	2.0	898	2	T01503
13	9	2.0	1363	2	T43220
14	9	2.0	1390	2	T30346
15	9	2.0	2391	2	T18410
16	8	1.8	101	2	I18220
17	8	1.8	108	2	I18212
18	8	1.8	157	2	S09814
19	8	1.8	158	1	T05710
20	8	1.8	162	1	B83130
21	8	1.8	211	2	D69413
22	8	1.8	229	1	PQB852
23	8	1.8	229	2	D90831
24	8	1.8	229	2	G85688
25	8	1.8	237	2	T47271
26	8	1.8	250	2	T47611
27	8	1.8	260	2	G84645
28	8	1.8	266	2	A34599
29	8	1.8	266	2	A34599

30	8	1.8	294	1	A37818	osteopontin precu
31	8	1.8	298	2	T52117	zinc finger protei
32	8	1.8	301	2	JC5811	osteopontin - rat
33	8	1.8	317	1	A25917	conserved hypochet
34	8	1.8	318	2	B69026	hypothetical prote
35	8	1.8	328	2	T16747	ribosomal protein
36	8	1.8	352	1	R5HSL0	50S ribosomal prot
37	8	1.8	352	2	G84266	protein PIN19.7 (1
38	8	1.8	368	2	G96668	hypothetical prote
39	8	1.8	374	2	T33173	hypothetical prote
40	8	1.8	379	2	T04645	protein kinase hom
41	8	1.8	390	2	T01451	hypothetical prote
42	8	1.8	391	2	D64366	hypothetical prote
43	8	1.8	422	2	S52578	hypothetical prote
44	8	1.8	426	2	T05676	serine/threonine-s
45	8	1.8	434	2	B54843	hypothetical prote
46	8	1.8	436	2	T51402	nemo, form II - fr
47	8	1.8	439	2	T31734	serine/threonine-s
48	8	1.8	471	2	T39232	hypothetical prote
49	8	1.8	477	2	A54843	probable serine th
50	8	1.8	482	2	S37845	nemo, form I - fru
51	8	1.8	485	2	T49237	transcription init
52	8	1.8	494	2	D84860	hypothetical prote
53	8	1.8	504	2	T10558	probable protein k
54	8	1.8	513	2	B96524	hypothetical prote
55	8	1.8	526	2	T47786	hypothetical prote
56	8	1.8	538	2	C83284	probable biotin-de
57	8	1.8	541	2	A12238	glutamine-fructose
58	8	1.8	560	2	T49839	hypothetical prote
59	8	1.8	567	2	JC5957	transforming growt
60	8	1.8	579	2	JC5955	transforming growt
61	8	1.8	584	2	T49206	phosphoinositide-s
62	8	1.8	590	2	S57594	hypothetical prote
63	8	1.8	606	2	JC5956	transforming growt
64	8	1.8	684	2	C96596	hypothetical prote
65	8	1.8	690	2	C96572	protein F12M16.4 (
66	8	1.8	746	2	G02838	enhancer-of-zeste
67	8	1.8	800	2	T02852	probable membrane
68	8	1.8	836	2	B96716	probable serine/th
69	8	1.8	840	2	S48975	hypothetical prote
70	8	1.8	847	1	A53800	mixed-linase prot
71	8	1.8	872	2	T18861	probable peptide-a
72	8	1.8	938	2	A56731	chromatin assembly
73	8	1.8	943	2	B45082	neurotrophic recep
74	8	1.8	992	2	T05335	hypothetical prote
75	8	1.8	992	2	S49835	hypothetical prote
76	8	1.8	1019	2	T13039	tyrosine kinase re
77	8	1.8	1038	2	T02634	rep protein homolo
78	8	1.8	1085	2	S55352	IFN1 protein - yea
79	8	1.8	1128	2	G86266	hypothetical prote
80	8	1.8	1131	2	S22266	FUN30 protein - ye
81	8	1.8	1131	2	S62982	vacuolar protein V
82	8	1.8	1165	2	T00363	hypothetical prote
83	8	1.8	1234	2	S66835	probable membrane
84	8	1.8	1341	2	T24490	hypothetical prote
85	8	1.8	2251	2	T18501	hypothetical prote
86	8	1.8	3394	2	T74501	cell division cycl
87	7	1.5	51	2	S66729	abscisic acid-indu
88	7	1.5	54	2	S66725	conserved hypochet
89	7	1.5	54	2	A69154	hypothetical prote
90	7	1.5	71	2	F97131	transcription repr
91	7	1.5	95	2	T46086	probable small sec
92	7	1.5	101	1	RGECKK	hypothetical prote
93	7	1.5	107	2	T35523	protein-serine/thr
94	7	1.5	108	2	T16893	conserved hypochet
95	7	1.5	111	2	I18218	ig lambda chain V
96	7	1.5	114	2	D46516	conserved hypochet
97	7	1.5	116	2	AD2693	hypothetical prote
98	7	1.5	123	2	T49794	hypothetical prote
99	7	1.5	133	2	T48978	PTS system, fructo
100	7	1.5	133	2	AE1435	

## ALIGNMENTS

## RESULT 1

B87950 Protein F33E2.2 (imported) - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001

C/Accession: B87950

R/Anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog

A/Reference number: A75000; MUID:99069613; PMID:9851916

A/Note: see websites genome.mutl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.leg

A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A/Accession: B87950

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-848 &lt;STO&gt;

A/Cross-references: GB:chr\_I; PIDN:CA806544.1; PID:G3876653; GSPDB:GN00019; CESP:F33E2.2

A/Genetics:

A/Map position: 1

Query Match 2.4%; Score 11; DB 2; Length 848;

Best Local Similarity 100.0%; Pred. No. 0.044;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 GVLWEMLTRE 203

Db 235 GVLWEMLTRE 245

## RESULT 2

T20082 Hypothetical protein F33E2.2 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999

C/Accession: T20082; T21703

R/Lennard, N.

Submitted to the EMBL Data Library, April 1998

A/Reference number: Z19221

A/Accession: T20082

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-855 &lt;WIL&gt;

A/Cross-references: EMBL:AL022593; PIDN:CAA18635.2; GSPDB:GN00019; CESP:F33E2.2

A/Experimental source: clone C49G9

A/Accession: T21703

A/Reference number: Z19461

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-855 &lt;WIL&gt;

A/Cross-references: EMBL:Z84574; PIDN:CA806544.2; GSPDB:GN00019; CESP:F33E2.2

A/Experimental source: clone F33E2

A/Accession: T21703

A/Reference number: Z19461

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-855 &lt;WIL&gt;

A/Cross-references: EMBL:Z84574; PIDN:CA806544.2; GSPDB:GN00019; CESP:F33E2.2

A/Experimental source: clone F33E2

A/Accession: T21703

A/Reference number: Z19461

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-855 &lt;WIL&gt;

A/Cross-references: EMBL:Z84574; PIDN:CA806544.2; GSPDB:GN00019; CESP:F33E2.2

A/Experimental source: clone F33E2

A/Accession: T21703

A/Reference number: Z19461

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-855 &lt;WIL&gt;

A/Cross-references: EMBL:Z84574; PIDN:CA806544.2; GSPDB:GN00019; CESP:F33E2.2

A/Experimental source: clone F33E2

KEX1 protein precursor - yeast (Saccharomyces cerevisiae)

N/Alternate names: protein G1237; protein YGL203C

C/Species: Saccharomyces cerevisiae

C/Date: 08-Mar-1989 #sequence\_revision 08-Mar-1989 #text\_change 21-Jul-2000

C/Accession: A29651; S64221

R/Dmochowska, A.; Dignard, D.; Henning, D.; Thomas, D.Y.; Bussey, H.

Cell 50, 573-584, 1987

A/Title: Yeast KEX1 gene encodes a putative protease with a carboxypeptidase B-like fun

A/Reference number: A29651; MUID:87273520; PMID:3301004

A/Accession: A29651

A/Molecule type: DNA

A/Residues: 1-729 &lt;DMO&gt;

A/Cross-references: GB:M17231; NID:G171778; PIDN:AAA4717.1; PID:G171779

R/Bjournson, A.J.; McReynolds, A.D.K.; Wright, L.F.

Submitted to the Protein Sequence Database, May 1996

A/Reference number: S64218

A/Accession: S64221

A/Molecule type: DNA

A/Residues: 1-729 &lt;BJO&gt;

A/Cross-references: EMBL:Z72725; NID:G1322835; PID:e243802; PID:G1322836; MIPS:YGL203C

A/Experimental source: strain S288C

C/Genetics:

A/Genes: SGD:KEX1

A/Cross-references: SGD:S0003171; MIPS:YGL203C

A/Map position: 7L

C/Keywords: Golgi apparatus; transmembrane protein

F1-22/Domain: signal sequence #status predicted &lt;SIG&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;

F1-22/Domain: signal sequence #status predicted &lt;MNT&gt;





A/Accession: T08069  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-408 <KIR>  
 A/Cross-references: EMBL:U06196; NID:g1019886; PIDN:AAA06956.1; PID:g1019887  
 A/Experimental source: strain 219T(+)  
 A/Function:  
 A/Description: may be involved in a signalling pathway in fertilization  
 A/Keywords: autophosphorylation

Query Match 2.0%; Score 9; DB 2; Length 408;  
 Best Local Similarity 100.0%; Pred. No. 2.1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 435 EGDGDDDD 443  
 DB 393 EGDGDDDD 401

## RESULT 8

C4337  
 Activin receptor like kinase-1 homolog - mouse

N/Alternate names: activin type IB receptor  
 C/Species: Mus musculus (house mouse)  
 C/Date: 06-Dec-1995 #sequence\_revision 08-Feb-1996 #text\_change 21-Jul-2000  
 C/Accession: J04337; I49478  
 R/Wu, X.; Robinson, C.E.; Fong, H.W.; Crabtree, J.S.; Rodriguez, B.R.; Roe, B.A.; Gimble  
 Biochem. Biophys. Res. Commun. 216, 78-83, 1995  
 A/Title: Cloning and characterization of the murine activin receptor like kinase-1 (ALK-1)  
 A/Reference number: J04337; MUID:96067535; PMID:7486127

A/Accession: J04337  
 A/Molecule type: mRNA  
 A/Residues: 1-502 <MX>  
 A/Cross-references: GB:I48015; NID:g1408066; PIDN:AAA03642.1; PID:g1020393  
 A/Experimental source: Lung  
 R/Dewulf, N.; Verschuere, K.; Lomoy, O.; Moren, A.; Grimsby, S.; Vande Spiegle, K.; M  
 Endocrinology 136, 2652-2663, 1995  
 A/Title: Distinct spatial and temporal expression patterns of two type I receptors for B  
 A/Reference number: I48241; MUID:95269711; PMID:7750489

A/Accession: I49478  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-502 <RES>  
 A/Cross-references: GB:I48015; NID:g1408066; PIDN:AAA03642.1; PID:g1020393  
 C/Comment: This protein belongs to transforming growth factor beta family, and it is a c

C/Genetics:  
 A/Gene: ALK-1  
 C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
 C/Keywords: ATP; glycoprotein; receptor; transmembrane protein  
 F:118-139/Domain: protein kinase homology <KIN>  
 F:199-495/Domain: protein kinase homology <KIN>  
 F:207-215/Region: protein kinase ATP-binding motif  
 F:197/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 2.0%; Score 9; DB 2; Length 502;  
 Best Local Similarity 100.0%; Pred. No. 2.5;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 131 HRDLKSRNV 139  
 DB 327 HRDLKSRNV 335

## RESULT 9

I48241

ALK-1 - mouse

C/Species: Mus musculus (house mouse)

C/Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 24-Sep-1999

C/Accession: I48241

R/Dewulf, N.; Verschuere, K.; Lomoy, O.; Moren, A.; Grimsby, S.; Vande Spiegle, K.; M

Endocrinology 136, 2652-2663, 1995

A/Title: Distinct spatial and temporal expression patterns of two type I receptors for B

A/Reference number: I48241; MUID:95269711; PMID:7750489

A/Accession: I48241  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-502 <RES>  
 A/Cross-references: EMBL:Z31664; NID:g840814; PIDN:CAA83484.1; PID:g840815  
 C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
 C/Keywords: ATP  
 F:199-495/Domain: protein kinase homology <KIN>  
 F:207-215/Region: protein kinase ATP-binding motif

Query Match 2.0%; Score 9; DB 2; Length 502;  
 Best Local Similarity 100.0%; Pred. No. 2.5;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 131 HRDLKSRNV 139  
 DB 327 HRDLKSRNV 335

## RESULT 10

S64951  
 Hypothetical protein YLR114C - Yeast (Saccharomyces cerevisiae)

N/Alternate names: hypothetical protein L2941

C/Species: Saccharomyces cerevisiae

C/Date: 01-Aug-1995 #sequence\_revision 24-May-1996 #text\_change 19-Apr-2002

C/Accession: S64951; S69401

R/Vernaselt, P.; Voet, M.; Volckaert, G.

Submitted to the Protein Sequence Database, May 1996

A/Reference number: S64943

A/Accession: S64951

A/Molecule type: DNA

A/Residues: 1-764 <VER>

A/Cross-references: EMBL:Z73286; NID:g1360509; PIDN:CAA97681.1; PID:e245807; PID:g13605

A/Experimental source: strain S288C

R/Vernaselt, P.; Volckaert, G.

Submitted to the EMBL Data Library, September 1995

A/Reference number: S69393

A/Accession: S69401

A/Molecule type: DNA

A/Residues: 1-764 <VER>

A/Cross-references: EMBL:X89514; NID:g1297019; PIDN:CAA61692.1; PID:e198747; PID:g12970

C/Genetics:

A/Gene: SGD:ERR4

A/Cross-references: SGD:S0004104

A/Map position: 12R

Query Match 2.0%; Score 9; DB 2; Length 764;  
 Best Local Similarity 100.0%; Pred. No. 3.5;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDGE 445  
 DB 628 DDDDDDDGE 636

## RESULT 11

TS2415

Polycarb protein E2A1 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #text\_change 24-Oct-2000

C/Accession: TS2415

R/Bilodeau, P.; Luo, M.; Dennis, E.S.; Peacock, W.J.; Chaudhury, A.M.

submitted to the EMBL Data Library, October 1998

A/Reference number: Z26069

A/Accession: TS2415

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-856 <BIL>

A/Cross-references: EMBL:AF100163; PIDN:AA09108.1

Query Match 2.0%; Score 9; DB 2; Length 856;  
 Best Local Similarity 100.0%; Pred. No. 3.9;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

	Number	of	Gaps	
y	172	WMAPEVQ	179	
b	92	WMAPEVQ	99	

## RESULT 17

protein-serine/threonine kinase - human (fragment)  
 C/Species: Homo sapiens (man)  
 C/Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 24-Sep-1999  
 C/Accession: I38212; S37414  
 R/Schultz, S.J.; Ni99, E.A.  
 Cell Growth Differ. 4, 821-830, 1993  
 A>Title: Identification of 21 novel human protein kinases, including 3 members of a family  
 A/Reference number: I38211; MUID:94100173; PMID:8274451  
 A/Accession: I38212  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-108 <RES>  
 A/Cross-references: EMBL:Z25422; NID:9405722; PIDN:CAA80909.1; PID:9405723  
 C/Suprafamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
 F:1-108/Domain: protein kinase homology (fragment) <KIN>

Query Match 1.8%; Score 8; DB 2; Length 108;  
 Best Local Similarity 100.0%; Pred. No. 6.3;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 172 WMAPEVIO 179  
 DB 92 WMAPEVIO 99

## RESULT 18

hypothetical protein UL51 - human cytomegalovirus (strain AD169)  
 C/Species: human cytomegalovirus, human herpesvirus 5  
 A/Note: host Homo sapiens (man)  
 C/Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 08-Oct-1999  
 C/Accession: S09814  
 R/Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Hornsneil, T.; M.; Barrell, B.G.  
 Curr. Top. Microbiol. Immunol. 154, 125-169, 1990  
 A/Title: Analysis of the protein-coding region of the sequence of human cytomegalovirus  
 A/Reference number: S09749; MUID:90269039; PMID:2161319  
 A/Accession: S09814  
 A/Status: nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-157 <CHE>  
 A/Cross-references: EMBL:X17403; NID:959591; PIDN:CAA35410.1; PID:91780829  
 A/Note: this sequence was submitted to the EMBL Data Library, December 1989  
 C/Suprafamily: varicella-zoster virus gene 25 protein

Query Match 1.8%; Score 8; DB 2; Length 157;  
 Best Local Similarity 100.0%; Pred. No. 8.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 440 DDDGEEE 447  
 DB 12 DDDGEEE 19

## RESULT 19

2S albumin precursor - soybean  
 N/Alternate names: aspartic acid-rich peptide  
 C/Species: Glycine max (soybean)  
 C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C/Accession: T05710; A28485  
 R/Wang, J.; Pichersky, E.  
 Plant Physiol. 114, 1567, 1997  
 A/Title: Nucleotide sequence of S-adenosyl-L-methionine:Caffeic Acid 3-O-Methyltransferase  
 A/Reference number: Z15424  
 A/Accession: T05710  
 A/Status: translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-158 <MAN>

A/Cross-references: EMBL:AF005030; NID:92305019; PID:92305020  
 A/Experimental source: Cultivar Hodgson 78; cotyledon

R/Odani, S.; Koide, T.; Ono, T.  
 J. Biol. Chem. 267, 10502-10505, 1997  
 A/Title: Amino acid sequence of a soybean (Glycine max) seed polypeptide having a poly(I  
 A/Reference number: A28485; MUID:87280104; PMID:3611081  
 A/Accession: A28485  
 A/Molecule type: protein  
 A/Residues: 22-64 <ODA>  
 A/Status: preliminary  
 C/Suprafamily: soybean 2S albumin  
 C/Keywords: glycoprotein  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-64/Product: 2S albumin small chain #status predicted <MAT>  
 F:54-56/Region: cell attachment (R-G-D) motif  
 F:39/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 1.8%; Score 8; DB 1; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 8.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDD 443  
 DB 55 GDDDDDD 62

## RESULT 20

hypothetical protein PA4129 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C/Species: Pseudomonas aeruginosa  
 C/Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C/Accession: B83130  
 R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; B. adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lablitz, K.; Lim, N.; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path  
 A/Reference number: A82950; MUID:20437337; PMID:10984043  
 A/Accession: B83130  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-162 <STO>  
 A/Cross-references: GB:AE004829; GB:AE004091; NID:99950327; PIDN:AA07516.1; GSPDB:GN00  
 A/Experimental source: strain PA01  
 A/Genetics:

Query Match 1.8%; Score 8; DB 2; Length 162;  
 Best Local Similarity 100.0%; Pred. No. 8.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 329 PLPLPLA 336  
 DB 34 PLPLPLA 41

## RESULT 21

hypothetical protein T2015.12 [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C/Accession: A96520  
 R/Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K  
 anen, N.F.; Hughes, B.; Hultzar, L.  
 Nature 408, 816-820, 2000  
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maitl, R.; Marziani  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,  
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Smith, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.  
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A/Reference number: A86141; MUID:21016719; PMID:11130712  
 A/Accession: A96520

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-198 <STO>  
A:Cross-references: GB:AE005173; NID:g10645475; PIDN:AAG21589.1; GSPDB:GN00141  
A:Gene: T2J15.12  
A:Map position: 1

Query Match  
Best Local Similarity 1.8%; Score 8; DB 2; Length 198;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
DB 101 GDDDDDD 108

## RESULT 22

Conserved hypothetical protein AF1309 - *Archaeoglobus fulgidus*  
A:Species: *Archaeoglobus fulgidus*  
A:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
C:Accession: D69413

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
Glocke, A.; Zhou, L.; Overbeek, R.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
A:Authors: Overbeek, R.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.;  
Smith, H.O.; Moore, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
A:Reference number: A69250; MUID:98049343; PMID:9389475  
A:Accession: D69413  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-211 <KLE>  
A:Cross-references: GB:AE001013; GB:AE000782; NID:g2689336; PIDN:AAB89938.1; PID:g264927

Query Match  
Best Local Similarity 1.8%; Score 8; DB 2; Length 211;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 PLLLLPLAA 336  
DB 174 PLLLLPLAA 181

## RESULT 23

QBP82  
A:Alternator Q - phase 82

A:Alternate names: regulatory protein Q  
C:Species: phase 82  
C:Date: 30-Sep-1990 #sequence\_revision 30-Sep-1990 #text\_change 23-Jul-1999  
R:Goiliger, J.A.; Roberts, J.W.  
A:Title: Bacteriophage 82 gene Q and O protein: sequence, overproduction, and activity  
A:Reference number: A29791; MUID:87308148; PMID:3624233  
A:Accession: A29791  
A:Molecule type: DNA  
A:Residues: 1-229 <GOI>

A:Cross-references: GB:J02803; NID:g215364; PIDN:AAA2298.1; PID:g215365  
A:Note: the authors translated the codon ACR for residue 188 as Phe and GAA for residue  
J. Mol. Biol. 257, 561-573, 1996  
A:Title: Holliday junction resolvases encoded by homologous rnaA genes in *Escherichia* cd

A:Reference number: S66579; MUID:96196428; PMID:8648624  
A:Accession: S66585  
A:Molecule type: DNA  
A:Residues: 1-229 <MAH>

A:Cross-references: EMBL:X92586; NID:g1051111; PIDN:CAA63332.1; PID:g1051118  
A:Gene: Q  
C:Superfamily: phase 82 regulatory protein Q

C:Keywords: DNA binding; late protein; transcription regulation

Query Match  
Best Local Similarity 1.8%; Score 8; DB 1; Length 229;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 TLERLKL 297  
DB 142 TLERLKL 149

## RESULT 24

antitermination protein [imported] - *Escherichia coli* (strain O157:H7, substrain RMD 0)  
A:Species: *Escherichia coli*  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
C:Accession: D90831

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gen  
A:Reference number: A99629; MUID:21156231; PMID:11236796  
A:Accession: D90831  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-229 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA835043.1; PID:g13161084; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RMD 050952  
A:Gene: EcolE20  
C:Superfamily: phase 82 regulatory protein Q

Query Match  
Best Local Similarity 1.8%; Score 8; DB 2; Length 229;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 TLERLKL 297  
DB 142 TLERLKL 149

## RESULT 25

probable antiterminator Q of prophage CP-933X 21874 [imported] - *Escherichia coli* (stra  
A:Species: *Escherichia coli*  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: G85688

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayh  
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; DiMaio, E.; Potamousis, K.; Apodaca  
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: G85688

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-229 <STO>  
A:Cross-references: GB:AE005174; NID:g12514802; PIDN:AA655971.1; GSPDB:GN00145; UWGP:21  
A:Experimental source: strain O157:H7, substrain EDL933  
A:Gene: 21874  
C:Superfamily: phase 82 regulatory protein Q

Query Match  
Best Local Similarity 1.8%; Score 8; DB 2; Length 229;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 TLERLKL 297  
DB 142 TLERLKL 149

## RESULT 26

T47271

hypothetical protein tmpr [imported] - Mycobacterium smegmatis insertion sequence IS1096  
 C:Species: Mycobacterium smegmatis  
 C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 21-Jul-2000  
 C:Accession: T47271  
 R:Carillo, J.D.; Barletta, R.G.; Bloom, B.R.; Jacobs, W.R.  
 J. Bacteriol. 173, 7772-7780, 1991  
 A:Title: A novel transposon trap for mycobacteria: isolation and characterization of IS1  
 A:Reference number: Z24439; MUID:92078082; PMID:1660454  
 A:Accession: T47271  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-237 <CIR>  
 A:Cross-references: EMBL:M76495; NID:G150003; PIDN:AAA98488.1; PID:G150004  
 A:Experimental source: insertion sequence IS1096  
 A:Genetics:  
 A:Gene: tmpr  
 A:Mobile element: insertion sequence IS1096

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 2; Length 237;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 FDLSEGD 438  
 |||||  
 Db 119 FDLSEGD 126

RESULT 27  
 T47611  
 hypothetical protein T14E10.90 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 28-Jul-2000  
 R:Obenaider, B.; Ottenwaelder, B.; Duchemin, D.; Zeidler, K.; Mewes, H.W.; Lemcke, K.; W  
 submitted to the Protein Sequence Database, March 2000  
 A:Reference number: Z24463  
 A:Accession: T47611  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-250 <OBE>  
 A:Cross-references: EMBL:A138656  
 A:Experimental source: cultivar Columbia; BAC clone T14E10  
 A:Genetics:  
 A:Map position: 3  
 A:Insertions: 73/3; 129/3; 188/3  
 A:Note: T14E10.90  
 C:Superfamily: Arabidopsis thaliana hypothetical protein T14E10.90

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 2; Length 250;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 RUKKLERD 300  
 |||||  
 Db 54 RUKKLERD 61

RESULT 28  
 G84645  
 hypothetical protein At2G25220 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: G84645  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKen, S.E.; Umayam, L.; Tallon, L.  
 enus, D.; Niernann, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: G84645  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-260 <STO>

A:Cross-references: GB:AE002093; NID:G4567255; PIDN:AAD23669.1; GSPDB:GN00139  
 A:Genetics:  
 A:Gene: At2G25220  
 A:Map position: 2

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 2; Length 260;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VHRDLKS 136  
 |||||  
 Db 125 VHRDLKS 132

RESULT 29  
 A34599  
 DNA-binding protein MCM1 - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: DNA-binding protein FUN80; DNA-binding protein PRRF/GRM; protein YW9;  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 31-Mar-2000  
 C:Accession: A34599; S14888; S22855; S52892  
 R:Ammerer, G.  
 Genes Dev. 4, 299-312, 1990  
 A:Title: Identification, purification, and cloning of a polypeptide (PRRF/GRM) that bin  
 A:Reference number: A34599; MUID:90249735; PMID:2159934  
 A:Accession: A34599  
 A:Molecule type: DNA  
 A:Residues: 1-286 <AMM>  
 A:Cross-references: GB:X52453; NID:G5266; PIDN:CAA36691.1; PID:G5267  
 R:Passmore, S.; Maine, G.T.; Elble, R.; Christ, C.; Tye, B.K.  
 J. Mol. Biol. 204, 593-606, 1988  
 A:Title: Saccharomyces cerevisiae protein involved in plasmid maintenance is necessary  
 A:Reference number: S14888; MUID:89141759; PMID:3066908  
 A:Accession: S14888  
 A:Molecule type: DNA  
 A:Residues: 1-286 <PAS>  
 A:Cross-references: EMBL:X14187; NID:G3909; PIDN:CAA32389.1; PID:G3910  
 R:Dubois, E.; Bercy, J.; Descamps, F.; Messenguy, F.  
 Gene 55, 265-275, 1987  
 A:Title: Characterization of two new genes essential for vegetative growth in Saccharom  
 A:Reference number: S22855; MUID:88030692; PMID:3311883  
 A:Accession: S22855  
 A:Molecule type: DNA  
 A:Residues: 1-8, 'S', '10-36, 'F', '38-155, 'AR' <DUB>  
 A:Cross-references: EMBL:M7511; NID:G171524; PIDN:AAA34609.1; PID:G171525  
 A:Note: the authors translated the codon TCT for residue 9 as Pro and TTC for residue 3  
 R:Odell, C.; Bowman, S.  
 submitted to the EMBL Data Library, February 1995  
 A:Reference number: S52885  
 A:Accession: S52885  
 A:Molecule type: DNA  
 A:Residues: 1-286 <OBE>  
 A:Cross-references: EMBL:Z48502; NID:G695715; PIDN:CAA8409.1; PID:G695723; MIPS:YMR043.  
 A:Genetics:  
 A:Gene: SGD:MCM1  
 A:Cross-references: SGD:S0004646; MIPS:YMR043w  
 A:Map position: 13R  
 C:Superfamily: transcription factor equa; serum response factor DNA-binding domain homo  
 C:Keywords: DNA binding; nucleus; transcription regulation  
 F:17-72/Domain: serum response factor DNA-binding domain homology <SRF>

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 2; Length 286;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
 |||||  
 Db 108 GDDDDDD 115

RESULT 30  
 A37818  
 osteopontin precursor - mouse

N/Alternate names: bone sialoprotein I; early T lymphocyte activation 1 protein; phospho  
 C/Species: Mus musculus (house mouse)  
 C/Date: 12-Feb-1993 #sequence revision 27-Jun-1994 #text change 10-Sep-1999  
 A/Accession: A37818; S04078; S12064; A33853; J10105; A60531; S11677  
 R/Miyazaki, Y.; Setoguchi, M.; Yoshida, S.; Higuchi, Y.; Akizuki, S.; Yamamoto, S.  
 J. Biol. Chem. 265, 14432-14438, 1990  
 A/Title: The mouse osteopontin gene. Expression in monocytic lineages and complete nucle  
 A/Reference number: A37818; PMID:90354433; PMID:2387863  
 A/Accession: A37818  
 A/Molecule type: DNA  
 A/Residues: 1-294 <MI>  
 A/Cross-references: EMBL:X51834  
 R/Miyazaki, Y.; Setoguchi, M.; Yoshida, S.; Higuchi, Y.; Akizuki, S.; Yamamoto, S.  
 Nucleic Acids Res. 17, 3298, 1989  
 A/Title: Nucleotide sequence of cDNA for mouse osteopontin-like protein.  
 A/Reference number: S04078; PMID:89263742; PMID:2726465  
 A/Accession: S04078  
 A/Molecule type: mRNA  
 A/Residues: 1-294 <MI2>  
 A/Cross-references: EMBL:X13986; NID:953755; PIDN:CAA32165.1; PID:953756  
 Yamamoto, S.  
 Submitted to the EMBL Data Library, January 1990  
 A/Reference number: S12064  
 A/Accession: S12064  
 A/Molecule type: DNA  
 A/Residues: 1-121, 'F', 123-294 <YAM>  
 A/Cross-references: EMBL:X51834; NID:953520; PIDN:CAA3132.1; PID:9297546  
 J. Biol. Chem. 264, 9682-9689, 1989  
 A/Title: Osteopontin, a transmembrane-associated cell adhesion phosphoprotein, is induc  
 A/Reference number: A33853; PMID:89255479; PMID:2722855  
 A/Accession: A33853  
 A/Molecule type: mRNA  
 A/Residues: 1-98, 'G', 100-294 <CRA>  
 A/Cross-references: GB:J04606; NID:9200157; PIDN:AA57265.1; PID:9200158  
 J. Pept. Med. 170, 145-161, 1989  
 A/Title: Structural and functional studies of the early T lymphocyte activation 1 (Eta-1)  
 on.  
 A/Reference number: J10105; PMID:89310352; PMID:2787378  
 A/Accession: J10105  
 A/Molecule type: mRNA  
 A/Residues: 1-42, 'P', 44-294 <PAT>  
 A/Cross-references: EMBL:X16151; NID:950863; PIDN:CAA34276.1; PID:950864  
 R/Singh, R.P.; Patarca, R.; Schwartz, J.; Singh, P.; Cantor, H.  
 J. Exp. Med. 171, 1931-1942, 1990  
 A/Title: Definition of a specific interaction between the early T lymphocyte activation  
 A/Reference number: A60931; PMID:90278349; PMID:2351930  
 A/Accession: A60931  
 A/Molecule type: protein  
 A/Residues: 158-176 <SIN>  
 C/Comment: This protein is an acidic glycoprotein rich in aspartic acid, glutamic acid,  
 C/Genetics:  
 A/Gene: Eta-1  
 A/Map position: 5  
 A/Introns: 18/3; 30/3; 57/3; 71/3; 165/3  
 C/Superfamily: osteopontin  
 C/Keywords: bone; cell binding; extracellular matrix; phosphoprotein; sialoglycoprotein  
 F:1-16/Doman: signal sequence #status predicted <SIG>  
 F:17-294/Product: osteopontin #status predicted <SIG>  
 F:85-96/Region: aspartic acid-rich  
 F:144-146/Region: cell attachment (R-G-D) motif  
 F:78/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 1.8%; Score 8; DB 1; Length 294;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
 |||||  
 DB 88 DDDDDDDG 95

RESULT 31  
 T52117  
 zinc finger protein [imported] - Arabidopsis thaliana (fragment)  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 20-Oct-2000 #sequence revision 20-Oct-2000 #text change 20-Oct-2000  
 A/Accession: T52117  
 R/Biodau, P.  
 Submitted to the EMBL Data Library, October 1997  
 A/Reference number: Z25963  
 A/Accession: T52117  
 A/Status: preliminary; translated from GB/EMBL/DBD  
 A/Molecule type: mRNA  
 A/Residues: 1-298 <BIL>  
 A/Cross-references: EMBL:AF030304; PIDN:AAC09174.1

Query Match 1.8%; Score 8; DB 2; Length 298;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
 |||||  
 DB 25 DDDDDDDG 32

RESULT 32  
 JCS811  
 osteopontin - rat  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 04-Feb-1998 #sequence revision 13-Mar-1998 #text change 07-May-1999  
 A/Accession: JCS811  
 R/Lasa, M.; Chang, P.L.; Prince, C.W.; Pina, L.A.  
 Biochem. Biophys. Res. Commun. 240, 602-605, 1997  
 A/Title: Phosphorylation of osteopontin by Golgi apparatus casein kinase.  
 A/Reference number: JCS811; PMID:98063283; PMID:9398611  
 A/Accession: JCS811  
 A/Molecule type: protein  
 A/Residues: 1-301 <LAS>  
 A/Experimental source: brain  
 C/Comment: This protein is involved in the initiation of the bone calcification process,  
 C/Superfamily: osteopontin

Query Match 1.8%; Score 8; DB 2; Length 301;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
 |||||  
 DB 72 DDDDDDDG 79

RESULT 33  
 A25917  
 osteopontin precursor - rat  
 N/Alternate names: bone sialoprotein I; phosphoprotein I, secreted; tumor-secreted phos  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 18-Dec-1997 #sequence revision 27-Jun-1994 #text change 10-Sep-1999  
 A/Accession: A25917; A45132; B45132; S28772; S04506; A45925  
 R/Olberg, A.; Franzén, A.; Heinegård, D.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 8819-8823, 1986  
 A/Title: Cloning and sequence analysis of rat bone sialoprotein (osteopontin) cDNA reve  
 A/Reference number: A25917; PMID:87067405; PMID:3024151  
 A/Accession: A25917  
 A/Molecule type: mRNA  
 A/Residues: 1-317 <OLD>  
 A/Cross-references: GB:M14656; NID:9205859; PIDN:AAA11762.1; PID:9205860  
 R/Singh, K.; Mukherjee, A.B.; De Vouge, M.W.; Mukherjee, B.B.  
 J. Biol. Chem. 267, 23847-23851, 1992  
 A/Title: Differential processing of osteopontin transcripts in rat kidney- and osteobla  
 A/Reference number: A45132; PMID:93054745; PMID:1429723  
 A/Accession: A45132  
 A/Molecule type: protein  
 A/Residues: 36-51 <SINI>

A:Experimental source: kidney  
 A:Note: sequence extracted from NCBI backbone (NCBIP:118869)  
 A:Accession: BA5132  
 A:Molecule type: protein  
 A:Residues: 277-282 <SIN2>  
 R:Pinne: sequence extracted from NCBI backbone (NCBIP:118871)  
 A:Pinne: C.W., Oosawa, T.; Butler, W.T.; Tomana, M.; Brown, A.S.; Bhown, M.; Schrotehl  
 J. Biol. Chem. 262, 2300-2307, 1987  
 A>Title: Isolation, characterization, and biosynthesis of a phosphorylated glycoprotein  
 A:Reference number: S28772; MUID:87137349; PMID:3469201  
 A:Accession: S28772  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 17-26 'X' <PRI>  
 R:Seinger, D.R.; Petruzzzi, C.A.; Papadopoulos, A.; Tenen, D.G.  
 Biochim. Biophys. Acta 966, 43-48, 1989  
 A>Title: Purification of a human milk protein closely similar to tumor-secreted phosphop  
 A:Reference number: S04505; MUID:89287357; PMID:2736258  
 A:Accession: S04506  
 A:Status: preliminary  
 A:Molecule type: protein  
 A:Residues: 17-26,155-167 <SEN>  
 R:Seinger, D.R.; Petruzzzi, C.A.; Gracey, C.F.; Papadopoulos, A.; Tenen, D.G.  
 Cancer Res. 48, 5770-5774, 1988  
 A>Title: Secreted phosphoproteins associated with neoplastic transformation: close homol  
 A:Reference number: A45925; MUID:89002730; PMID:3167835  
 A:Accession: A45925  
 A:Molecule type: protein  
 A:Residues: 17-25 <S22>  
 C:Comment: this protein is an acidic glycoprotein rich in aspartic acid, glutamic acid,  
 C:superfamily: osteopontin  
 C:keywords: bone; cell binding; extracellular matrix; phosphoprotein; sialoglycoprotein  
 F:1-16/Domain: signal sequence #status predicted <SIG>  
 F:17-317/Product: osteopontin #status predicted <MAT>  
 F:86-96/Region: aspartic acid-rich  
 F:144-146/Region: cell attachment (R-G-D) motif  
 F:79/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.8%; Score 8; DB 1; Length 317;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
 |||||  
 88 DDDDDDDG 95

RESULT 34  
 9026  
 reserved hypothetical protein MTH1196 - Methanobacterium thermoautotrophicum (strain De  
 C:Species: Methanobacterium thermoautotrophicum  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #ext\_change 26-May-2000  
 C:Accession: E69026  
 R:Smith, D.R.; Donette-Stamm, L.A.; DeLoughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
 R:Ou, D.; Spadefora, R.; Viciore, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.  
 K.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
 J. Bacteriol. 179, 7133-7155, 1997  
 A>Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
 A:Reference number: A69000; MUID:98037514; PMID:9371143  
 A:Accession: E69026  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-318 <MTH>  
 A:Cross-references: GB:AE000888; GB:AE000666; NID:92622304; PIDN:AB85685.1; PID:9262230  
 A:Experimental source: strain Delta H  
 C:Genetics:  
 A:Gene: MTH1196  
 A:Start codon: TTG  
 C:superfamily: conserved hypothetical protein MTH1196

Query Match 1.8%; Score 8; DB 2; Length 318;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

44 VKKLIKIE 51
| | | | |
Db 28 VKKLIKIE 35

RESULT 35
T16747
Hypotheical protein R13F6.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000
C:Accession: T16747
R:Miller, N.
submitted to the EMBL Data Library, April 1994
A:Description: The sequence of C. elegans cosmid R13F6.
A:Reference number: Z18570
A:Accession: T16747
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-328 <M1>
A:Cross-References: EMBL:U00046; NID:9470358; PID:9470364; PID:MAC47047.1; GSPDB:GN000.
A:Experimental source: strain Bristol N2; clone R13F6
C:Genetics:
A:Gene: CESP:R13F6.7
A:Map position: 3
A:Intons: 20/3; 160/3; 222/2; 286/2
C:Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 i
Query Match 1.8%; Score 8; DB 2; Length 328;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 EKEAEILS 58
| | | | |
Db 90 EKEAEILS 97

RESULT 36
RSHSL0
Ribosomal protein L10 [similarity] - Halobacterium salinarum
N:Alternate names: ribosomal protein P0
C:Species: Halobacterium salinarum
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 21-Jul-2000
C:Accession: S01315; S04120; S11587
R:Joch, T.
Eur. J. Biochem. 176, 297-303, 1988
A>Title: Complete nucleotide sequence of the ribosomal 'A' protein operon from the arch
A:Reference number: S01314; MUID:88329082; PMID:2458258
A:Accession: S01315
A:Molecule type: DNA
A:Residues: 1-352 <IIO>
A:Cross-References: EMBL:X13008; NID:943532; PID:CAA31431.1; PID:943534
A:Experimental source: strain S9
A>Note: The source is designated as Halobacterium halobium
R:Shimmin, L.C.; Dennis, P.P.
EMBO J. 8, 1223-1235, 1989
A>Title: Characterization of the L11, L1, L10 and L12 equivalent ribosomal protein genes
A:Reference number: S04116; MUID:89305527; PMID:2743981
A:Accession: S04120
A:Molecule type: DNA
A:Residues: 1-58, 'V', 60-352 <SH1>
A:Cross-References: EMBL:X15078; NID:943449; PID:CAA33180.1; PID:943454
A>Note: The source is designated as Halobacterium cutirubrum
C:Superfamily: rat acidic ribosomal protein P0
C:Keywords: protein biosynthesis; ribosome

Query Match 1.8%; Score 8; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
| | | | |
Db 333 DDDDDDDG 340

```



## RESULT 37

G84266  
50S ribosomal protein L10P [imported] - Halobacterium sp. NRC-1  
C/Species: Halobacterium sp. NRC-1  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C/Accession: G84266  
RING, W.V.; Kennedy, S.P.; Mahatras, G.G.; Bernquist, B.; Pan, M.; Shukla, H.D.; Laeky, S.; Lelchauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabic, Junc. K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li, A.  
A/Title: Genome sequence of Halobacterium species NRC-1.  
A/Reference number: A84160; MUID:20504483; PMID:11016950  
A/Accession: G84266  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-352 <STO>  
A/Cross-references: GB:AE004437; NID:G10580651; PIDN:AAQ19499.1; GSPDB:GN00138  
C/Genetics:  
A/Map position:  
C/Superfamily: rat acidic ribosomal protein P0

Query Match 1.8%; Score 8; DB 2; Length 352;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 437 DDDDDDDG 444  
DB 333 DDDDDDDG 340

## RESULT 38

G96668  
protein F1N19.7 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 01-Mar-2002  
C/Accession: G96668  
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso, J.; Holt, T.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marziani, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Selzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, R.; M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: G96668  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-368 <STO>  
A/Cross-references: GB:AE005173; NID:G6633811; PIDN:AAF19670.1; GSPDB:GN00141  
C/Genetics:  
A/Map position:  
C/Superfamily: Arabidopsis thaliana hypothetical protein T4P9.90

Query Match 1.8%; Score 8; DB 2; Length 368;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 437 DDDDDDDG 444  
DB 61 DDDDDDDG 68

## RESULT 39

T33173  
hypothetical protein C24G6.5 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C/Accession: T33173  
R/Greco, T.; Bradshaw, H.; Keppler, D.

submitted to the EMBL Data Library, May 1998  
A/Description: The sequence of C. elegans cosmid C24G6.  
A/Reference number: 221298  
A/Accession: T33173  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-374 <GRE>  
A/Cross-references: EMBL:AF067936; PIDN:AAQ19208.1; GSPDB:GN00023; CESP:C24G6.5  
A/Experimental source: strain Bristol N2; clone C24G6  
C/Genetics:  
A/Map position:  
A/Introns: 83/1; 351/3

Query Match 1.8%; Score 8; DB 2; Length 374;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 436 GDDDDDD 443  
DB 189 GDDDDDD 196

## RESULT 40

T04645  
hypothetical protein F10N7.190 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 20-Sep-1999  
C/Accession: T04645  
R/Bevan, M.; Rose, M.; Hempel, S.; Ertlan, K.D.; Hohnle, J.; Mewes, H.W.; Mayer, K.F.; submitted to the Protein Sequence Database, March 1999  
A/Reference number: 215263  
A/Accession: T04645  
A/Molecule type: DNA  
A/Residues: 1-379 <BEV>  
A/Cross-references: EMBL:AL021636  
C/Genetics:  
A/Map position:  
A/Introns: 48/1; 221/1; 252/3; 287/1; 326/3  
A/Note: F10N7.190  
C/Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 1.8%; Score 8; DB 2; Length 379;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 129 VTHRDLS 136  
DB 233 VTHRDLS 240

## RESULT 41

T01451  
protein kinase homolog F2401.13 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 23-Mar-2001  
C/Accession: T01451  
R/Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Coe, J.; Ecker, J.R.  
submitted to the EMBL Data Library, January 1998  
A/Description: Genomic sequence for Arabidopsis thaliana BAC F2401.  
A/Reference number: Z14211  
A/Accession: T01451  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-390 <SHI>  
A/Cross-references: EMBL:AC003113; NID:G2689438; PID:G2781357; GSPDB:GN00059; ATSP:F2401

C/Genetics:  
A/Map position:  
A/Introns: 83/1; 351/3

A:Map position: 1  
A:Introns: 149/3; 301/3  
C:Superfamily: Kinase-related transforming protein; protein kinase homology

Query Match 1.8%; Score 8; DB 2; Length 390;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VVHRDLKS 136  
|||||  
DB 208 VVHRDLKS 215

## RESULT 42

D64366  
hypothetical protein homolog M0532 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000

C:Accession: D64366  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Bult, C.J.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodex, A.; Reich, C.I.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurett, M.A.; Jon, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurett, M.A.

Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A:Reference number: A64300; MUID:96337999; PMID:8688087

A:Accession: D64366  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA  
A:Residues: 1-391 <BUL>

A:Cross-references: GB:U67502; GB:L77117; NID:G2826293; PIDN:AAB98523.1; PID:91591235; T

C:Genetics:  
A:Map position: REV468785-467610

C:Superfamily: hypothetical protein M0532

Query Match 1.8%; Score 8; DB 2; Length 391;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 VVIMADGV 146  
|||||  
DB 149 VVIMADGV 156

RESULT 43  
S52578  
serine/threonine-specific protein kinase NPK15 (EC 2.7.1.-) - common tobacco

C:Species: Nicotiana tabacum (common tobacco)  
C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 24-Sep-1999

C:Accession: S52578  
Mol. Gen. Genet. 245, 1-10, 1994

A:Title: NPK15, a tobacco protein-kinase with a single hydrophobic reg-

A:Reference number: S52578; MUID:95147840; PMID:7845351

A:Accession: S52578  
A:Molecule type: mRNA

A:Residues: 1-422 <ITO>

A:Cross-references: EMBL:D31737, NID:9505145, PIDN:BA06538.1; PID:9505146

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol

C:Keywords: ATP, phosphotransferase; serine/threonine-specific protein kinase

F:114-382/Domain: protein kinase homology <KIN>

F:122-130/Region: protein kinase ATP-binding motif

Query Match 1.8%; Score 8; DB 2; Length 422;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VVHRDLKS 136  
|||||  
DB 235 VVHRDLKS 242

RESULT 44

T05676  
hypothetical protein F20M13.40 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 23-Jul-1999

C:Accession: T05676  
R:Bevan, M.; Wedler, H.; Kutzner, M.; Wambolt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.I.

submitted to the Protein Sequence Database, February 1999

A:Reference number: Z15420

A:Accession: T05676

A:Molecule type: DNA

A:Residues: 1-426 <BEV>

A:Cross-references: EMBL:AL035540

A:Experimental source: cultivar Columbia; BAC clone F20M13

C:Genetics:  
A:Map position: 4

A:Introns: 90/3; 149/2; 204/1; 350/3

A>Note: F20M13.40

Query Match 1.8%; Score 8; DB 2; Length 426;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
|||||  
DB 407 GDDDDDD 414

RESULT 45  
B54843  
nemato, form II - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change 19-Dec-1997

C:Accession: B54843

R:Choi, K.W.; Benzer, S.

Cell 78, 125-136, 1994

A:Title: Rotation of photoreceptor clusters in the developing Drosophila eye requires ti

A:Reference number: A54843; MUID:94306509; PMID:8033204

A:Accession: B54843

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-434 <CHO>

A:Cross-references: GB:U12009

C:Genetics:  
A:Gene: nmc

A:Cross-references: FlyBase:Fgn011817

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol

C:Keywords: alternative splicing; ATP

F:38-301/Domain: protein kinase homology <KIN>

F:46-54/Region: protein kinase ATP-binding motif

Query Match 1.8%; Score 8; DB 2; Length 434;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VVKICDPG 153  
|||||  
DB 179 VVKICDPG 166

RESULT 46  
T51402  
serine/threonine-specific protein kinase-like protein - Arabidopsis thaliana

N:Alternate names: protein F14P8.110

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 17-Nov-2000

C:Accession: T51402

R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Me

submitted to the Protein Sequence Database, August 2000

A:Reference number: Z25394

A:Accession: T51402

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-436 <SAT>

A:Cross-references: EMBL:AL391144  
A:Experimental source: cultivar Columbia; BAC clone F14F8  
C:Genetics:  
A:Map position: 5  
A:Introns: 107/2; 158/3; 202/1; 254/3; 322/3  
A:Note: F14F8 110  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

Query Match 1.8%; Score 8; DB 2; Length 436;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 129 VTHRDLS 136  
DB 235 VTHRDLS 242

## RESULT 47

biochemical protein T05C3.5 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 31-Jan-2000  
C:Accession: T31734  
R:Blanchard, M.; Bradshaw, H.; Kramer, J.  
Submitted to the EMBL Data Library, July 1997  
A:Description: The sequence of C. elegans cosmid T05C3.  
A:Reference number: 221076  
A:Accession: T31734  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-439 <BLA>  
A:Cross-references: EMBL:AF016428; PIDN:AB65361.1; GSPDB:GN00023; CESP:T05C3.5  
C:Genetics:  
A:Gene: CESP:T05C3.5  
A:Map position: 5  
A:Introns: 267/3; 298/2; 377/3  
C:Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology

Query Match 1.8%; Score 8; DB 2; Length 439;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 436 GDDDDDD 443  
DB 105 GDDDDDD 112

## RESULT 48

probable serine threonine protein kinase - fission yeast (*Schizosaccharomyces pombe*)  
C:Species: *Schizosaccharomyces pombe*  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
C:Accession: T39232  
R:Churcher, C.M.; Gentles, S.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
Submitted to the EMBL Data Library, August 1997  
A:Reference number: 221837  
A:Accession: T39232  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-471 <CHU>  
A:Cross-references: EMBL:Z98763; PIDN:CA1193.1; GSPDB:GN00066; SPDB:SPAC9G1.09  
C:Genetics:  
A:Experimental source: strain 972h-; cosmid c9G1  
A:Gene: SPDB:SPAC9G1.09  
A:Map position: 1  
A:Introns: 39/3  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

Query Match 1.8%; Score 8; DB 2; Length 471;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179  
DB 171 WMAPEVIQ 178

## RESULT 49

A54843  
memo, form I - fruit fly (*Drosophila melanogaster*)  
C:Species: *Drosophila melanogaster*  
C:Date: 03-Oct-1995 #sequence\_revision 03-Oct-1995 #text\_change 24-Sep-1999  
C:Accession: A54843  
R:Choi, K.W.; Benzer, S.  
Cell 78, 125-136, 1994  
A:Title: Rotation of photoreceptor clusters in the developing *Drosophila* eye requires C.  
A:Reference number: A54843; MUID:94306509; PMID:8033204  
A:Accession: A54843  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-477 <CHO>  
A:Cross-references: GB:U12009; NID:9515669; PIDN:AAA21124.1; PID:9532558  
C:Genetics:  
A:Gene: nmo  
A:Cross-references: FlyBase:FBgn0011817  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C:Keywords: alternative splicing; ATP  
F:38-301/Domain: protein kinase homology <KIN>  
F:46-54/Region: protein kinase ATP-binding motif

Query Match 1.8%; Score 8; DB 2; Length 477;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 146 VLKICDFG 153  
DB 179 VLKICDFG 186

## RESULT 50

S37845  
transcription initiation factor IIE chain TPA1 - yeast (*Saccharomyces cerevisiae*)  
N:Alternate names: protein YKL028w; transcription initiation factor a 66k chain  
C:Species: *Saccharomyces cerevisiae*  
A:Variety: strain S288C  
C:Date: 23-Feb-1995 #sequence\_revision 11-Aug-1995 #text\_change 21-Jul-2000  
C:Accession: S37845; A55465; C55465  
R:Rieger, M.  
Submitted to the Protein Sequence Database, March 1994  
A:Reference number: S37832  
A:Accession: S37845  
A:Molecule type: DNA  
A:Residues: 1-482 <RIE>  
A:Cross-references: EMBL:Z28028; NID:9486026; PID:9486027; MIPS:YKL028w  
J:Feaver, W.J.; Henry, N.L.; Bushnell, D.A.; Sayre, M.H.; Brickner, J.H.; Gileadi, O.;  
J. Biol. Chem. 269, 27549-27553, 1994  
A:Title: Yeast TFIIE. Cloning, expression, and homology to vertebrate proteins.  
A:Reference number: A55465; MUID:95050500; PMID:7961670  
A:Accession: A55465  
A:Molecule type: DNA  
A:Residues: 1-482 <FEA>  
A:Cross-references: GB:U12825; NID:9607957; PIDN:AAA62665.1; PID:9607958  
A:Accession: C55465  
A:Molecule type: protein  
A:Residues: 102-108; 206-220 <FE2>  
C:Genetics:  
A:Gene: SGD:TPA1  
A:Cross-references: SGD:S0001511; MIPS:YKL028w  
A:Map position: 11L  
C:Keywords: nucleus; transcription initiation

Query Match 1.8%; Score 8; DB 2; Length 482;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
DB 464 GDDDDDD 471

## RESULT 51

T49237

hypothetical protein F7K15.80 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000

C:Accession: T49237

R:Obermayer, B.; Oltmannseder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Rudd, S.; Lem

submitted to the Protein Sequence Database, April 2000

A:Reference number: 225019

A:Accession: T49237

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-485 &lt;OBE&gt;

A:Cross-references: EMBL:AL353871; GSPDB:GN00061; ATSP:F7K15.80

A:Experimental source: cultivar Columbia; BAC clone F7K15

A:Gene: ATSP:F7K15.80

A:Map position: 3

A:introns: 292/3; 359/3; 425/1

Query Match

Best Local Similarity 100.0%; Pred. No. 23;

Matches 8; Conservative 0; Mismatches 0; Gaps 0;

QY 437 DDDDDDDG 444

DB 25 DDDDDDDG 32

RESULT 52

D84860

probable protein kinase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: D84860

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Xoo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.

euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: D84860

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-494 &lt;STO&gt;

A:Cross-references: GB:AE002093; NID:94512659; PIDN:AMD21713.1; GSPDB:GN00139

C:Genetics:

A:Gene: AT2G42960

A:Map position: 2

Query Match

Best Local Similarity 100.0%; Pred. No. 23;

Matches 8; Conservative 0; Mismatches 0; Gaps 0;

QY 41 EVAVKKL 48

DB 207 EVAVKKL 214

RESULT 53

T10558

hypothetical protein T12G13.150 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 15-Oct-1999

C:Accession: T10558

R:Bevan, M.; Lemard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro

submitted to the Protein Sequence Database, June 1999

A:Reference number: Z16533

Query Match

Best Local Similarity 100.0%; Pred. No. 23;

Matches 8; Conservative 0; Mismatches 0; Gaps 0;

QY 41 EVAVKKL 48

DB 207 EVAVKKL 214

RESULT 54

B96524

hypothetical protein F1A17.5 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001

C:Accession: B96524

R:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzalli,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzbey, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: B96524

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-513 &lt;STO&gt;

A:Cross-references: GB:AE005173; NID:95733869; PIDN:AMD49757.1; GSPDB:GN00141

C:Genetics:

A:Gene: F1A17.5

A:Map position: 1

C:Superfamily: Arabidopsis thaliana hypothetical protein F1J16.30

Query Match

Best Local Similarity 100.0%; Pred. No. 24;

Matches 8; Conservative 0; Mismatches 0; Gaps 0;

QY 437 DDDDDDDG 444

DB 319 DDDDDDDG 326

RESULT 55

T47786

hypothetical protein F1J16.90 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T47786

R:D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Rudd, S.

submitted to the Protein Sequence Database, April 2000

A:Reference number: Z24476

A:Accession: T47786

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-526 &lt;DAN&gt;

A:Cross-references: EMBL:AL163527

A:Experimental source: cultivar Columbia; BAC clone F1J16

C:Genetics:

A:Map position: 3

A:introns: 67/3; 247/3; 289/2; 322/2

A/Note: .f17J16.90

Query Match  
Best Local Similarity 1.8%; Score 8; DB 2; Length 526;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 EGGDDDD 442  
|||||  
DB 491 EGGDDDD 498

## RESULT 56

C83284

probable biotin-dependent carboxylase PA2888 [Imported] - Pseudomonas aeruginosa (strain  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: C83284

R:Stover, C.K.; Pam, X.O.; Eryn, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.  
Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Lardig, K.; Lam,  
Lory, S.; Olson, M.V.  
C:Map position: 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: C83284  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-538 <STO>  
A:Cross-references: GB:AE004715; GB:AE004091; NID:99948977; PIDN:AAG06276.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:

A:Gene: PA2888

C:Superfamily: propionyl-CoA carboxylase beta chain

Query Match  
Best Local Similarity 1.8%; Score 8; DB 2; Length 538;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 324 EGSNTPL 331  
|||||  
DB 364 EGSNTPL 371

## RESULT 57

A12238

glutamine-fructose-6-phosphatetransferase [Imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp.

A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
A:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
A:Accession: A12238

R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi,  
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 203-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: A12238  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-541 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BA075163.1; PID:G17132597; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:

A:Gene: nodM

C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)

Query Match  
Best Local Similarity 1.8%; Score 8; DB 2; Length 541;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 286 EIEATLR 293  
|||||  
DB 375 EIEATLR 382

## RESULT 58

T49839

hypothetical protein B24H17.250 [Imported] - Neurospora crassa  
C:Species: Neurospora crassa  
C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
C:Accession: T49839

R:Schulte, U.; Algen, V.; Hohnsbeil, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura  
submitted to the Protein Sequence Database, May 2000  
A:Reference number: Z25022  
A:Accession: T49839  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-560 <SCH>  
A:Cross-references: EMBL:AJ356815; GSPDB:GN00116; NCSP:B24H17.250  
A:Experimental source: BAC clone B24H17; strain OR74A  
C:Genetics:

A:Gene: NCSP:B24H17.250

A:Map position: 6

A:Introns: 72/1; 142/3; 270/1; 310/1; 380/3; 411/2; 446/2; 468/1; 485/3

Query Match  
Best Local Similarity 1.8%; Score 8; DB 2; Length 560;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
|||||  
DB 385 DDDDDDDG 392

## RESULT 59

JC5957

transforming growth factor-beta activated kinase (EC 2.7.-.-) 1c - human  
C:Species: Homo sapiens (man)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 17-Mar-2000  
C:Accession: JC5957

R:Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.  
Biochem. Biophys. Res. Commun. 243, 545-549, 1998  
A:Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-in  
A:Reference number: JC5955; MUID:98153801; PMID:9480845  
A:Accession: JC5957  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-567 <SAK>  
A:Cross-references: DDBJ:AB009358  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol.

C:Keywords: phosphotransferase

Query Match  
Best Local Similarity 1.8%; Score 8; DB 2; Length 567;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKIDFG 153  
|||||  
DB 170 VLKIDFG 177

## RESULT 60

JC5955

transforming growth factor-beta activated kinase (EC 2.7.-.-) 1a - human  
C:Species: Homo sapiens (man)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: JC5955

R:Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.  
Biochem. Biophys. Res. Commun. 243, 545-549, 1998  
A:Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-in  
A:Reference number: JC5955; MUID:98153801; PMID:9480845  
A:Accession: JC5955  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-579 <SAK>  
A:Cross-references: DDBJ:AB009356; NID:92924623; PIDN:BA025025.1; PID:G2924624  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol.

C:Keywords: phosphotransferase

Query Match 1.8%; Score 8; DB 2; Length 579;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153  
 DB 170 VLKICDFG 177

## RESULT 61

T49206 phosphoinositide-specific phospholipase C-like protein - Arabidopsis thaliana  
 N:Alternate names: protein F27K19.120  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 28-Jun-2000  
 C:Accession: T49206  
 R:Benes, V.; Wurmbach, E.; Drzonek, H.; Ansoerge, W.; Mewes, H.W.; Rudd, S.; Lemcke, K.;  
 submitted to the Protein Sequence Database, April 2000  
 A:Reference number: Z25014  
 A:Accession: T49206  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-584 <BEN>  
 A:Cross-references: EMBL:AL163832; GSPDB:GN00061; ATSP:F27K19.120  
 A:Experimental source: cultivar Columbia; BAC clone F27K19  
 C:Genetics:  
 A:Gene: ATSP:F27K19.120  
 A:Map position: 3  
 A:Insertion: 100/3; 166/2; 211/3; 290/3; 371/2; 410/3; 458/3; 487/3  
 C:Superfamily: Arabidopsis thaliana 1-phosphatidylinositol 4,5-bisphosphate phospholipase  
 bisphosphate phospholipase domain Y homology  
 F:105-248/Domain: 1-phosphatidylinositol 4,5-bisphosphate phospholipase domain X hom

Query Match 1.8%; Score 8; DB 2; Length 584;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
 DB 295 GDDDDDD 302

## RESULT 62

S57594 hypothetical protein YMR27C - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: hypothetical protein YMR27C.1  
 C:Species: Saccharomyces cerevisiae  
 C>Date: 19-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 29-Oct-1995  
 C:Accession: S57594  
 R:Skellton, J.; Churcher, C.M.;  
 submitted to the EMBL Data Library, June 1995  
 A:Reference number: S57587  
 A:Accession: S57594  
 A:Molecule type: DNA  
 A:Residues: 1-590 <SK>  
 A:Cross-references: EMBL:Z49939; NID:9887599; PIDN:CAA90198.1; PID:9887608; MIPS:YMR27C  
 A:Experimental source: strain AB972  
 C:Genetics:  
 A:Gene: SGD:TAF67  
 A:Cross-references: SGD:S0004840; MIPS:YMR27C  
 A:Map position: 13R

Query Match 1.8%; Score 8; DB 2; Length 590;  
 Best Local Similarity 100.0%; Pred. No. 27;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
 DB 413 DDDDDDDG 420

## RESULT 63

JC5956 transforming growth factor-beta activated kinase (EC 2.7.1.1) 1b - human  
 C:Species: Homo sapiens (man)  
 C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
 C:Accession: JC5956  
 R:Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.  
 Biochem. Biophys. Res. Commun. 243, 545-549, 1998  
 A:Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-in  
 A:Reference number: JC5955; PMID:98153801; PMID:9480845  
 A:Accession: JC5956  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-606 <SAK>  
 A:Cross-references: DDBJ:AB009357; NID:92924625; PIDN:BA25026.1; PID:92924626  
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol.  
 C:Keywords: phosphotransferase

Query Match 1.8%; Score 8; DB 2; Length 606;  
 Best Local Similarity 100.0%; Pred. No. 27;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153  
 DB 170 VLKICDFG 177

## RESULT 64

C96596 hypothetical protein T1813.3 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: C96596  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.  
 Chin, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Matli, R.; Matzli,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; PMID:21016719; PMID:11130712  
 A:Accession: C96596  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-684 <STO>  
 A:Cross-references: GB:AE005173; NID:G11094789; PIDN:AA29721.1; GSPDB:GN00141  
 C:Genetics:  
 A:Gene: T1813.3  
 A:Map position: 1

Query Match 1.8%; Score 8; DB 2; Length 684;  
 Best Local Similarity 100.0%; Pred. No. 30;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
 DB 162 GDDDDDD 169

## RESULT 65

C96572 protein F12M16.4 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: C96572  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.  
 Chin, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Matli, R.; Matzli,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: G96572  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-690 <STO>  
 A:Cross-references: GB:AE005173; NID:97769851; PIDN:AAF69529.1; GSPDB:GN00141  
 A:Gene: F12M16.4  
 A:Map position: 1

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 2; Length 690;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179  
 187 WMAPEVIQ 194

RESULT 66  
 G02838  
 enhancer-of-zeste homolog 2 - human  
 C:Species: Homo sapiens (man)  
 C>Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 17-Jul-1998  
 R:Accession: G02838  
 R:Author(s): S.B.  
 submitted to the EMBL Data Library, February 1996  
 A:Reference number: H01746  
 A:Accession: G02838  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-746 <ANT>  
 A:Cross-references: EMBL:X95653; NID:G1438063  
 C:Gene(s):  
 A:Gene: GDB:EZH2; EZH1  
 A:Cross-references: GDB:701613; OMIM:601573  
 A:Map position: 21q22.2-21q22.2

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 2; Length 746;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
 183 DDDDDDDG 190

RESULT 67  
 T02852  
 Probable membrane protein L1439.4 [imported] - Leishmania major (strain Friedlin)  
 C:Species: Leishmania major  
 C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 19-May-2000  
 R:Accession: T02852; H81462  
 R:Author(s): P.J.  
 submitted to the EMBL Data Library, May 1998  
 A:Description: The nucleotide sequence of Leishmania major Friedlin chromosome 1.  
 A:Reference number: 214740  
 A:Accession: T02852  
 A>Status: preliminary  
 A:Molecule type: translated from GB/EMBL/DBJ  
 A:Residues: 1-800 <MTL>  
 A:Cross-references: EMBL:AE001274; NID:G3264850; PID:G2266920  
 R:Author(s): P.W.; Audleman, L.; DeVos, T.; Hixson, G.; Kleier, P.; Lemley, C.; Magness, C.;  
 Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999  
 A>Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c  
 A:Reference number: A81455; MUID:99178987; PMID:10077609  
 A:Accession: H81462  
 A>Status: preliminary  
 A:Molecule type: DNA

A:Residues: 1-800 <PVL>  
 A:Cross-references: GB:AE001274; NID:G3264850; PIDN:AAAC24675.1; PID:G2266920; GSPDB:GN0.  
 A:Experimental source: strain MHOM/IL/81/Friedlin  
 C:Gene(s):  
 A:Gene: L1439.4  
 A:Map position: 1

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 2; Length 800;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
 Db 21 DDDDDDDG 28

RESULT 68  
 B96716  
 probable serine/threonine kinase F23010.20 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 R:Accession: B96716  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Matli, R.; Marzella,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: B96716  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-836 <STO>  
 A:Cross-references: GB:AE005173; NID:97705100; PIDN:AAF67779.1; GSPDB:GN00141  
 C:Gene(s):  
 A:Gene: F23010.20  
 A:Map position: 1

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 2; Length 836;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179  
 Db 412 WMAPEVIQ 419

RESULT 69  
 S48975  
 Hypothetical protein YHR131c - yeast (Saccharomyces cerevisiae)  
 C:Species: Saccharomyces cerevisiae  
 C>Date: 02-Dec-1994 #sequence\_revision 02-Dec-1994 #text\_change 19-Apr-2002  
 R:Accession: S48975  
 R:Fulton, L.  
 submitted to the EMBL Data Library, June 1994  
 A:Description: The sequence of S. cerevisiae cosmid 9315.  
 A:Reference number: S48967  
 A:Accession: S48975  
 A:Molecule type: DNA  
 A:Residues: 1-840 <FUL>  
 A:Cross-references: EMBL:U10398; NID:G551328; PID:G500681; GSPDB:GN00008; MIPS:YHR131c  
 C:Gene(s):  
 A:Gene: MIPS:YHR131c  
 A:Cross-references: SGD:S0001173  
 A:Map position: 8R

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 2; Length 840;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
 |||||  
 Db 829 DDDDDDDG 836

## RESULT 70

mixed-lineage protein kinase (EC 2.7.1.-) 3 - human  
 A:Accession: A53800  
 N:Alternate names: protein kinase PTKI; protein kinase SPRK  
 C:Species: Homo sapiens (man)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: A53800; 158395  
 R:Gallo, K.A.; Mark, M.R.; Scadden, D.T.; Wang, Z.; Gu, Q.; Godowski, P.J.  
 J:Bio. Chem. 269, 15092-15100, 1994  
 A:Title: Identification and characterization of SPRK, a novel src-homology 3 domain-cont  
 A:Reference number: A53800; MUID:94253068; PMID:8195146  
 A:Accession: A53800  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-847 <GAL>  
 A:Cross-references: GB:U07747; NID:9464027; PIDN:AAA19647.1; PID:9464028  
 R:King, Y.L.; Leung, I.W.; Heng, H.H.; Tsui, L.C.; Lasseam, N.J.  
 Oncogene 9, 1745-1750, 1994  
 A:Title: MLK-3: identification of a widely-expressed protein kinase bearing an SH3 domain  
 A:Reference number: 158395; MUID:94239754; PMID:8183572  
 A:Accession: 158395  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-847 <RES>  
 A:Cross-references: GB:L32976; NID:9488295; PIDN:AAA59859.1; PID:9488296  
 A:Gene: GDB:MLK3; PTKI; SPRK  
 C:Cross-references: GDB:134755; OMIM:600050  
 A:Map position: 11q13.1-11q13.3  
 C:Superfamily: mixed-lineage protein kinase 3; protein kinase homology; SH3 homology  
 C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k

F:48-100/Domain: SH3 homology <SH32>  
 F:115-383/Domain: protein kinase homology <KIN>  
 F:123-131/Region: protein kinase ATP-binding motif  
 F:403-424/Region: leucine zipper motif  
 F:438-459/Region: leucine zipper motif  
 F:468-482/Region: basic

Query Match 1.8%; Score 8; DB 1; Length 847;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 129 VIHRLDKS 136  
 |||||  
 Db 237 VIHRLDKS 244

## RESULT 71

T18861  
 Probable peptidase-aspartate beta-dioxygenase (EC 1.14.11.16) - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
 C:Accession: T18861; 723533  
 R:Swaburne, J.  
 Submitted to the EMBL Data Library, August 1996  
 A:Reference number: Z19032  
 A:Accession: T18861  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-872 <WIL>  
 A:Cross-references: EMBL:Z79596; NID:e1323798; PIDN:CA01859.1; GSPDB:GN000028; CESP:K09A  
 R:Swaburne, J.  
 Submitted to the EMBL Data Library, August 1996  
 A:Reference number: Z19753  
 A:Accession: T23533  
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-872 <M12>  
 A:Cross-references: EMBL:Z79601; PIDN:CA01887.1; GSPDB:GN000028; CESP:K09A9.6  
 A:Experimental source: clone K09A9  
 C:Genetics: CESP:K09A9.6  
 A:Gene: CESP:K09A9.6  
 A:Map position: X  
 A:Interons: 18/1; 40/3; 74/1; 104/2; 139/1; 170/3; 385/3; 432/2; 464/3; 530/2; 609/3; 776

C:Keywords: oxidoreductase

Query Match 1.8%; Score 8; DB 2; Length 872;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
 |||||  
 Db 288 GDDDDDD 295

## RESULT 72

A56731  
 Chromatin assembly factor I p150 chain - human  
 C:Species: Homo sapiens (man)  
 C:Date: 08-Feb-1996 #sequence\_revision 08-Feb-1996 #text\_change 21-Jul-2000  
 C:Accession: A56731  
 R:Kauffman, P.D.; Kobayashi, R.; Kessler, N.; Stillman, B.  
 Cell 81, 1105-1114, 1995  
 A:Title: The p150 and p60 subunits of chromatin assembly factor I: a molecular link bet  
 A:Reference number: A56731; MUID:95323966; PMID:7600578  
 A:Accession: A56731  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-938 <KAU>  
 A:Cross-references: GB:U20979; NID:9882257; PIDN:AAA76736.1; PID:9882258

Query Match 1.8%; Score 8; DB 2; Length 938;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 434 SEGDDDD 441  
 |||||  
 Db 598 SEGDDDD 605

## RESULT 73

B45082  
 neurotrophic receptor tor2 precursor - human  
 N:Contains: protein-tyrosine kinase (EC 2.7.1.112)  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 19-May-2000  
 C:Accession: B45082  
 R:Maslakowski, P.; Carroll, R.D.  
 J: Biol. Chem. 267, 26181-26190, 1992  
 A:Title: A novel family of cell surface receptors with tyrosine kinase-like domain.  
 A:Reference number: A45082; MUID:93100347; PMID:1334494  
 A:Accession: B45082  
 A:Molecule type: mRNA  
 A:Residues: 1-943 <MAS>  
 A:Cross-references: GB:M97639; NID:9337466; PIDN:AAA60276.1; PID:9337467  
 A:Note: sequence extracted from NCBI backbone (NCBIP:120918)

C:Genetics: CESP:K09A9.6  
 A:Gene: GDB:NTRK2  
 A:Cross-references: GDB:136454  
 A:Map position: 6p21-6p21  
 C:Superfamily: neurotrophic receptor tor; immunoglobulin homology; kringle homology; pr  
 C:Keywords: ATP; glycoprotein; kringle; phosphotransferase; transmembrane protein; tyro  
 F:1-27/Domain: signal sequence #status predicted <SIG>  
 F:28-337/Product: neurotrophic receptor tor2 #status predicted <IMM>  
 F:76-137/Domain: immunoglobulin homology <IMM>  
 F:116-394/Domain: kringle homology <KR>  
 F:412-428/Domain: transmembrane #status predicted <TM>  
 F:471-753/Domain: protein kinase homology <KIN>  
 F:479-487/Region: protein kinase ATP-binding motif



F:70,188,318/Binding site: carbohydrate (Aen) (covalent) #status predicted

Query Match 1.8%; Score 8; DB 2; Length 943;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVLMW 198  
DB 677 SYGVLMW 684

## RESULT 74

T05335

hypothetical protein FIC12.190 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 20-Sep-1999

C/Accession: T05335  
R/Bevan, M.; Terry, N.; Ardley, W.; Buysheart, C.; Daseville, R.; De

mitted to the Protein Sequence Database, April 1998  
Reference number: 215408

A/Accession: T05335

A/Molecule type: DNA

A/Residues: 1-992 <BEV>

A/Cross-references: EMBL:AL022224

A/Experimental source: cultivar Columbia; BAC clone FIC12

C/Genetics:

A/Map position: 4

A/Introns: 884/1

A/Note: FIC12.190

C/Superfamily: protein kinase Xa21; leucine-rich alpha-2-glycoprotein repeat homology; F

## Query Match

1.8%; Score 8; DB 2; Length 992;  
Best Local Similarity 100.0%; Pred. No. 42;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 EVAVKKLL 48  
DB 734 EVAVKKLL 741

## RESULT 75

S49835

hypothetical protein YDR080W - yeast (Saccharomyces cerevisiae)

N/Alternate names: hypothetical protein D4446; hypothetical protein YD554.13

C/Species: Saccharomyces cerevisiae

C/Date: 13-Jan-1995 #sequence\_revision 10-Feb-1995 #text\_change 29-Oct-1999

C/Accession: S49835; S48767; S55827; S67897

R/Richard, C.; Harris, D.E.  
submitted to the EMBL Data Library, November 1994

A/Reference number: S49823

A/Accession: S49835

A/Molecule type: DNA

A/Residues: 1-992 <RIC>

A/Cross-references: EMBL:Z46796; NID:G577794; PIDN:CAA86802.1; PID:G577807

R/Coester, F.; Joniaux, J.L.; Goffeau, A.  
submitted to the EMBL Data Library, October 1994

A/Reference number: S48758

A/Accession: S48757

A/Molecule type: DNA

A/Residues: 1-423; 'M', 425-992 <COS>

A/Cross-references: EMBL:X82086; NID:G558241; PIDN:CAA57607.1; PID:G558251

R/Coester, F.; Joniaux, J.L.; Goffeau, A.

Yeast 11, 673-679, 1995

A/Title: Analysis of a 32.8 kb segment of yeast chromosome IV reveals 21 open reading fr

A/Reference number: S55819; PMID:96093910; PMID:7483840

A/Accession: S55827

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-423; 'M', 425-992 <COM>

A/Cross-references: EMBL:X82086; NID:G558241; PIDN:CAA57607.1; PID:G558251  
R/Poury, F.; Joniaux, J.L.; Fumelle, B.; Coester, F.; Goffeau, A.

submitted to the Protein Sequence Database, July 1996  
A/Reference number: S67889

A/Accession: S67897

A/Molecule type: DNA

A/Residues: 1-992 <FOU>

A/Cross-references: EMBL:Z74376; NID:G1431548; PIDN:CAA98899.1; PID:G253331; PID:G14315

A/Experimental source: strain S288C

C/Genetics:

A/Gene: SGD:VP541

A/Cross-references: SGD:S0002487; MIPS:YDR080W

A/Map position: 4R

## Query Match

1.8%; Score 8; DB 2; Length 992;  
Best Local Similarity 100.0%; Pred. No. 42;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
DB 78 GDDDDDD 85

Search completed: May 1, 2003, 20:56:24  
Job time : 73 secs



GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:39:47 ; Search time 26 Seconds

(without alignments)  
725,836 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 455

Sequence: 1 MSSLAGSPVQIKFDDLPFE.....GDDDDDDGDEEDNDNDNSE 455

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

d size: 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 100 summaries

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	2.2	729	1 KEX1_YEAST	P09620 saccharomyc
2	10	2.2	1089	1 KEX1_YEAST	P38798 saccharomyc
3	9	2.0	375	1 MCK1_YEAST	P21965 saccharomyc
4	9	2.0	502	1 KIR3_MOUSE	O61288 mus musculu
5	9	2.0	505	1 KIR3_RAT	P08203 ratu
6	9	2.0	1363	1 INSR_BRALA	O02466 branchisto
7	9	2.0	1390	1 INSR_BRALA	O02466 branchisto
8	8	1.8	157	1 ULS1_HCMVA	P16792 aedes aegypt
9	8	1.8	158	1 ZSS_SOYBN	P13594 glycin
10	8	1.8	229	1 REGO_BP82	P13870 bacterioph
11	8	1.8	286	1 MCM1_YEAST	P11746 saccharomyc
12	8	1.8	294	1 OSTP_MOUSE	P10923 mus musculu
13	8	1.8	317	1 OSTP_MOUSE	P10923 mus musculu
14	8	1.8	317	1 OSTP_MOUSE	P10923 mus musculu
15	8	1.8	352	1 RLA0_HALCU	P08721 ratu
16	8	1.8	352	1 RLA0_HALCU	P08721 ratu
17	8	1.8	391	1 Y532_METUA	P13553 halobacteri
18	8	1.8	482	1 Y532_METUA	P13553 halobacteri
19	8	1.8	487	1 Y532_METUA	P13553 halobacteri
20	8	1.8	491	1 Y532_METUA	P13553 halobacteri
21	8	1.8	579	1 M3K7_MOUSE	O05021 saccharomyc
22	8	1.8	590	1 M3K7_MOUSE	O05021 saccharomyc
23	8	1.8	606	1 M3K7_MOUSE	O05021 saccharomyc
24	8	1.8	746	1 EZH2_HUMAN	O61188 mus musculu
25	8	1.8	746	1 EZH2_HUMAN	O61188 mus musculu
26	8	1.8	799	1 CN3B_MOUSE	O61409 mus musculu
27	8	1.8	840	1 YHT1_YEAST	P38935 saccharomyc
28	8	1.8	875	1 DD10_HUMAN	O13206 homo sapien
29	8	1.8	938	1 CAPA_HUMAN	O13206 homo sapien
30	8	1.8	943	1 ROR2_HUMAN	O01974 homo sapien
31	8	1.8	944	1 ROR2_MOUSE	O01974 homo sapien
32	8	1.8	992	1 VP41_YEAST	O92138 mus musculu
33	8	1.8	1085	1 IFH1_YEAST	P38950 saccharomyc
			1131	1 YAB9_YEAST	P31380 saccharomyc

34	8	1.8	1165	1 YNF4_YEAST	P53950 saccharomyc
35	7	1.5	86	1 RNFH_AZOVI	O9F50 azobacter
36	7	1.5	101	1 KOA2_ECOLI	P03052 escherichia
37	7	1.5	147	1 UCR1_YEAST	P00127 saccharomyc
38	7	1.5	160	1 HPPK_HAETIN	P43777 haemophilus
39	7	1.5	172	1 UBC2_YEAST	P06104 saccharomyc
40	7	1.5	173	1 NUGM_BRABR	P43194 brachyramph
41	7	1.5	173	1 NUGM_BRABR	P43194 brachyramph
42	7	1.5	173	1 NUGM_BRABR	P43194 brachyramph
43	7	1.5	175	1 VIRR_AGRTU	O52278 agrobacteri
44	7	1.5	183	1 NO29_XENLA	O42584 xenopus lae
45	7	1.5	204	1 HMG1_ONCMY	O07713 mus musculu
46	7	1.5	207	1 HIT_MOUSE	O07713 mus musculu
47	7	1.5	228	1 EF1B_ORYSA	O04680 oryza sativ
48	7	1.5	230	1 EF1B_BETVU	O81918 beta vulgar
49	7	1.5	245	1 ICP3_HSV11	P37319 herpes simp
50	7	1.5	248	1 ICP3_HSV11	P37319 herpes simp
51	7	1.5	250	1 YB95_ARATH	O8487 arabidopsis
52	7	1.5	252	1 ICP3_HSV1D	P37318 herpes simp
53	7	1.5	253	1 ASP_PLAUS	P13825 plaesodiu
54	7	1.5	253	1 LI39_CAEBL	P34684 caenorhabd
55	7	1.5	253	1 T2D7_RAT	O62880 ratu
56	7	1.5	263	1 ICP3_HSV1F	P08353 herpes simp
57	7	1.5	264	1 OSTP_CHICK	P23498 gallus gall
58	7	1.5	264	1 T2D7_HUMAN	O16594 homo sapien
59	7	1.5	265	1 EF1D_XENLA	P29693 xenopus lae
60	7	1.5	280	1 HAX1_MOUSE	O35387 mus musculu
61	7	1.5	289	1 YNV9_CAEBL	P34682 caenorhabd
62	7	1.5	295	1 UBC3_YEAST	P14682 saccharomyc
63	7	1.5	303	1 PEX5_CANAL	O74711 candida alb
64	7	1.5	305	1 MURB_BACSU	P16579 bacillus su
65	7	1.5	309	1 CDK3_HUMAN	O00526 homo sapien
66	7	1.5	310	1 MRP_DROME	O9hkv0 drosophila
67	7	1.5	314	1 DOS2_YEAST	P54858 saccharomyc
68	7	1.5	314	1 MKO3_MOUSE	O63844 m. m. m. m. m.
69	7	1.5	323	1 KRAF_MSUV6	P00532 murine sarc
70	7	1.5	323	1 LUKF_STANU	P07726 streptoloco
71	7	1.5	325	1 HUGB_STANU	O02066 ctenophorab
72	7	1.5	328	1 KRAP_MOUSE	P28028 mus musculu
73	7	1.5	332	1 AARP_WHEAT	O02066 ctenophorab
74	7	1.5	332	1 CDG2_CAEBL	O02066 ctenophorab
75	7	1.5	333	1 Y713_CHLMU	O02066 ctenophorab
76	7	1.5	333	1 SP72_YEAST	O02066 ctenophorab
77	7	1.5	348	1 RLA0_HALMA	O9p1w3 chlamydia m
78	7	1.5	349	1 STY1_SCHPO	P15825 halosarcu
79	7	1.5	353	1 ASK2_ARATH	P15825 halosarcu
80	7	1.5	353	1 FUS3_YEAST	O09892 escherichia
81	7	1.5	355	1 ARG3_YEAST	P43292 arabidopsis
82	7	1.5	358	1 MKO1_MOUSE	P16892 saccharomyc
83	7	1.5	359	1 PKX1_HUMAN	P07250 saccharomyc
84	7	1.5	360	1 KMO5_XENLA	P27703 mus musculu
85	7	1.5	360	1 KMO5_XENLA	P27703 mus musculu
86	7	1.5	361	1 MKO1_HUMAN	P12965 xenopus lae
87	7	1.5	363	1 ASK1_ARATH	P28482 bos taurus
88	7	1.5	363	1 ASK1_ARATH	P28482 bos taurus
89	7	1.5	365	1 M14B_DROME	P26696 xenopus lae
90	7	1.5	366	1 M14B_DROME	P26696 xenopus lae
91	7	1.5	367	1 MK12_MOUSE	O61443 arabidopsis
92	7	1.5	367	1 MK12_MOUSE	O61443 arabidopsis
93	7	1.5	368	1 MK12_MOUSE	O61443 arabidopsis
94	7	1.5	370	1 CTFP_PLAUF	O08911 mus musculu
95	7	1.5	370	1 CTFP_PLAUF	O08911 mus musculu
96	7	1.5	370	1 CTFP_PLAUF	O08911 mus musculu
97	7	1.5	370	1 CTFP_PLAUF	O08911 mus musculu
98	7	1.5	371	1 NTF6_TOBAC	O63558 ratu
99	7	1.5	372	1 NTF6_TOBAC	O63558 ratu
100	7	1.5	373	1 YF04_YEAST	P43603 saccharomyc

RESULT 1

## ALIGNMENTS

KEX1\_YEAST  
 ID\_KEX1\_YEAST STANDARD: PRT; 729 AA.  
 AC P09630;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Carboxypeptidase KEX1 precursor (EC 3.4.16.6) (Carboxypeptidase D).  
 GN KEX1 OR YGL203C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OC NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87273520; PubMed=3301004;  
 RA Dmochowska A., Dignard D., Henning D., Thomas D.Y., Bussey H.;  
 RT "Yeast KEX1 gene encodes a putative protease with a carboxypeptidase  
 B-like function involved in killer toxin and alpha-factor precursor  
 processing.";  
 Cell 50:573-584 (1987).  
 [2]  
 [3]  
 [4]  
 [5]  
 [6]  
 [7]  
 [8]  
 [9]  
 [10]  
 [11]  
 [12]  
 [13]  
 [14]  
 [15]  
 [16]  
 [17]  
 [18]  
 [19]  
 [20]  
 [21]  
 [22]  
 [23]  
 [24]  
 [25]  
 [26]  
 [27]  
 [28]  
 [29]  
 [30]  
 [31]  
 [32]  
 [33]  
 [34]  
 [35]  
 [36]  
 [37]  
 [38]  
 [39]  
 [40]  
 [41]  
 [42]  
 [43]  
 [44]  
 [45]  
 [46]  
 [47]  
 [48]  
 [49]  
 [50]  
 [51]  
 [52]  
 [53]  
 [54]  
 [55]  
 [56]  
 [57]  
 [58]  
 [59]  
 [60]  
 [61]  
 [62]  
 [63]  
 [64]  
 [65]  
 [66]  
 [67]  
 [68]  
 [69]  
 [70]  
 [71]  
 [72]  
 [73]  
 [74]  
 [75]  
 [76]  
 [77]  
 [78]  
 [79]  
 [80]  
 [81]  
 [82]  
 [83]  
 [84]  
 [85]  
 [86]  
 [87]  
 [88]  
 [89]  
 [90]  
 [91]  
 [92]  
 [93]  
 [94]  
 [95]  
 [96]  
 [97]  
 [98]  
 [99]  
 [100]  
 [101]  
 [102]  
 [103]  
 [104]  
 [105]  
 [106]  
 [107]  
 [108]  
 [109]  
 [110]  
 [111]  
 [112]  
 [113]  
 [114]  
 [115]  
 [116]  
 [117]  
 [118]  
 [119]  
 [120]  
 [121]  
 [122]  
 [123]  
 [124]  
 [125]  
 [126]  
 [127]  
 [128]  
 [129]  
 [130]  
 [131]  
 [132]  
 [133]  
 [134]  
 [135]  
 [136]  
 [137]  
 [138]  
 [139]  
 [140]  
 [141]  
 [142]  
 [143]  
 [144]  
 [145]  
 [146]  
 [147]  
 [148]  
 [149]  
 [150]  
 [151]  
 [152]  
 [153]  
 [154]  
 [155]  
 [156]  
 [157]  
 [158]  
 [159]  
 [160]  
 [161]  
 [162]  
 [163]  
 [164]  
 [165]  
 [166]  
 [167]  
 [168]  
 [169]  
 [170]  
 [171]  
 [172]  
 [173]  
 [174]  
 [175]  
 [176]  
 [177]  
 [178]  
 [179]  
 [180]  
 [181]  
 [182]  
 [183]  
 [184]  
 [185]  
 [186]  
 [187]  
 [188]  
 [189]  
 [190]  
 [191]  
 [192]  
 [193]  
 [194]  
 [195]  
 [196]  
 [197]  
 [198]  
 [199]  
 [200]  
 [201]  
 [202]  
 [203]  
 [204]  
 [205]  
 [206]  
 [207]  
 [208]  
 [209]  
 [210]  
 [211]  
 [212]  
 [213]  
 [214]  
 [215]  
 [216]  
 [217]  
 [218]  
 [219]  
 [220]  
 [221]  
 [222]  
 [223]  
 [224]  
 [225]  
 [226]  
 [227]  
 [228]  
 [229]  
 [230]  
 [231]  
 [232]  
 [233]  
 [234]  
 [235]  
 [236]  
 [237]  
 [238]  
 [239]  
 [240]  
 [241]  
 [242]  
 [243]  
 [244]  
 [245]  
 [246]  
 [247]  
 [248]  
 [249]  
 [250]  
 [251]  
 [252]  
 [253]  
 [254]  
 [255]  
 [256]  
 [257]  
 [258]  
 [259]  
 [260]  
 [261]  
 [262]  
 [263]  
 [264]  
 [265]  
 [266]  
 [267]  
 [268]  
 [269]  
 [270]  
 [271]  
 [272]  
 [273]  
 [274]  
 [275]  
 [276]  
 [277]  
 [278]  
 [279]  
 [280]  
 [281]  
 [282]  
 [283]  
 [284]  
 [285]  
 [286]  
 [287]  
 [288]  
 [289]  
 [290]  
 [291]  
 [292]  
 [293]  
 [294]  
 [295]  
 [296]  
 [297]  
 [298]  
 [299]  
 [300]  
 [301]  
 [302]  
 [303]  
 [304]  
 [305]  
 [306]  
 [307]  
 [308]  
 [309]  
 [310]  
 [311]  
 [312]  
 [313]  
 [314]  
 [315]  
 [316]  
 [317]  
 [318]  
 [319]  
 [320]  
 [321]  
 [322]  
 [323]  
 [324]  
 [325]  
 [326]  
 [327]  
 [328]  
 [329]  
 [330]  
 [331]  
 [332]  
 [333]  
 [334]  
 [335]  
 [336]  
 [337]  
 [338]  
 [339]  
 [340]  
 [341]  
 [342]  
 [343]  
 [344]  
 [345]  
 [346]  
 [347]  
 [348]  
 [349]  
 [350]  
 [351]  
 [352]  
 [353]  
 [354]  
 [355]  
 [356]  
 [357]  
 [358]  
 [359]  
 [360]  
 [361]  
 [362]  
 [363]  
 [364]  
 [365]  
 [366]  
 [367]  
 [368]  
 [369]  
 [370]  
 [371]  
 [372]  
 [373]  
 [374]  
 [375]  
 [376]  
 [377]  
 [378]  
 [379]  
 [380]  
 [381]  
 [382]  
 [383]  
 [384]  
 [385]  
 [386]  
 [387]  
 [388]  
 [389]  
 [390]  
 [391]  
 [392]  
 [393]  
 [394]  
 [395]  
 [396]  
 [397]  
 [398]  
 [399]  
 [400]  
 [401]  
 [402]  
 [403]  
 [404]  
 [405]  
 [406]  
 [407]  
 [408]  
 [409]  
 [410]  
 [411]  
 [412]  
 [413]  
 [414]  
 [415]  
 [416]  
 [417]  
 [418]  
 [419]  
 [420]  
 [421]  
 [422]  
 [423]  
 [424]  
 [425]  
 [426]  
 [427]  
 [428]  
 [429]  
 [430]  
 [431]  
 [432]  
 [433]  
 [434]  
 [435]  
 [436]  
 [437]  
 [438]  
 [439]  
 [440]  
 [441]  
 [442]  
 [443]  
 [444]  
 [445]  
 [446]  
 [447]  
 [448]  
 [449]  
 [450]  
 [451]  
 [452]  
 [453]  
 [454]  
 [455]  
 [456]  
 [457]  
 [458]  
 [459]  
 [460]  
 [461]  
 [462]  
 [463]  
 [464]  
 [465]  
 [466]  
 [467]  
 [468]  
 [469]  
 [470]  
 [471]  
 [472]  
 [473]  
 [474]  
 [475]  
 [476]  
 [477]  
 [478]  
 [479]  
 [480]  
 [481]  
 [482]  
 [483]  
 [484]  
 [485]  
 [486]  
 [487]  
 [488]  
 [489]  
 [490]  
 [491]  
 [492]  
 [493]  
 [494]  
 [495]  
 [496]  
 [497]  
 [498]  
 [499]  
 [500]  
 [501]  
 [502]  
 [503]  
 [504]  
 [505]  
 [506]  
 [507]  
 [508]  
 [509]  
 [510]  
 [511]  
 [512]  
 [513]  
 [514]  
 [515]  
 [516]  
 [517]  
 [518]  
 [519]  
 [520]  
 [521]  
 [522]  
 [523]  
 [524]  
 [525]  
 [526]  
 [527]  
 [528]  
 [529]  
 [530]  
 [531]  
 [532]  
 [533]  
 [534]  
 [535]  
 [536]  
 [537]  
 [538]  
 [539]  
 [540]  
 [541]  
 [542]  
 [543]  
 [544]  
 [545]  
 [546]  
 [547]  
 [548]  
 [549]  
 [550]  
 [551]  
 [552]  
 [553]  
 [554]  
 [555]  
 [556]  
 [557]  
 [558]  
 [559]  
 [560]  
 [561]  
 [562]  
 [563]  
 [564]  
 [565]  
 [566]  
 [567]  
 [568]  
 [569]  
 [570]  
 [571]  
 [572]  
 [573]  
 [574]  
 [575]  
 [576]  
 [577]  
 [578]  
 [579]  
 [580]  
 [581]  
 [582]  
 [583]  
 [584]  
 [585]  
 [586]  
 [587]  
 [588]  
 [589]  
 [590]  
 [591]  
 [592]  
 [593]  
 [594]  
 [595]  
 [596]  
 [597]  
 [598]  
 [599]  
 [600]  
 [601]  
 [602]  
 [603]  
 [604]  
 [605]  
 [606]  
 [607]  
 [608]  
 [609]  
 [610]  
 [611]  
 [612]  
 [613]  
 [614]  
 [615]  
 [616]  
 [617]  
 [618]  
 [619]  
 [620]  
 [621]  
 [622]  
 [623]  
 [624]  
 [625]  
 [626]  
 [627]  
 [628]  
 [629]  
 [630]  
 [631]  
 [632]  
 [633]  
 [634]  
 [635]  
 [636]  
 [637]  
 [638]  
 [639]  
 [640]  
 [641]  
 [642]  
 [643]  
 [644]  
 [645]  
 [646]  
 [647]  
 [648]  
 [649]  
 [650]  
 [651]  
 [652]  
 [653]  
 [654]  
 [655]  
 [656]  
 [657]  
 [658]  
 [659]  
 [660]  
 [661]  
 [662]  
 [663]  
 [664]  
 [665]  
 [666]  
 [667]  
 [668]  
 [669]  
 [670]  
 [671]  
 [672]  
 [673]  
 [674]  
 [675]  
 [676]  
 [677]  
 [678]  
 [679]  
 [680]  
 [681]  
 [682]  
 [683]  
 [684]  
 [685]  
 [686]  
 [687]  
 [688]  
 [689]  
 [690]  
 [691]  
 [692]  
 [693]  
 [694]  
 [695]  
 [696]  
 [697]  
 [698]  
 [699]  
 [700]  
 [701]  
 [702]  
 [703]  
 [704]  
 [705]  
 [706]  
 [707]  
 [708]  
 [709]  
 [710]  
 [711]  
 [712]  
 [713]  
 [714]  
 [715]  
 [716]  
 [717]  
 [718]  
 [719]  
 [720]  
 [721]  
 [722]  
 [723]  
 [724]  
 [725]  
 [726]  
 [727]  
 [728]  
 [729]  
 [730]  
 [731]  
 [732]  
 [733]  
 [734]  
 [735]  
 [736]  
 [737]  
 [738]  
 [739]  
 [740]  
 [741]  
 [742]  
 [743]  
 [744]  
 [745]  
 [746]  
 [747]  
 [748]  
 [749]  
 [750]  
 [751]  
 [752]  
 [753]  
 [754]  
 [755]  
 [756]  
 [757]  
 [758]  
 [759]  
 [760]  
 [761]  
 [762]  
 [763]  
 [764]  
 [765]  
 [766]  
 [767]  
 [768]  
 [769]  
 [770]  
 [771]  
 [772]  
 [773]  
 [774]  
 [775]  
 [776]  
 [777]  
 [778]  
 [779]  
 [780]  
 [781]  
 [782]  
 [783]  
 [784]  
 [785]  
 [786]  
 [787]  
 [788]  
 [789]  
 [790]  
 [791]  
 [792]  
 [793]  
 [794]  
 [795]  
 [796]  
 [797]  
 [798]  
 [799]  
 [800]  
 [801]  
 [802]  
 [803]  
 [804]  
 [805]  
 [806]  
 [807]  
 [808]  
 [809]  
 [810]  
 [811]  
 [812]  
 [813]  
 [814]  
 [815]  
 [816]  
 [817]  
 [818]  
 [819]  
 [820]  
 [821]  
 [822]  
 [823]  
 [824]  
 [825]  
 [826]  
 [827]  
 [828]  
 [829]  
 [830]  
 [831]  
 [832]  
 [833]  
 [834]  
 [835]  
 [836]  
 [837]  
 [838]  
 [839]  
 [840]  
 [841]  
 [842]  
 [843]  
 [844]  
 [845]  
 [846]  
 [847]  
 [848]  
 [849]  
 [850]  
 [851]  
 [852]  
 [853]  
 [854]  
 [855]  
 [856]  
 [857]  
 [858]  
 [859]  
 [860]  
 [861]  
 [862]  
 [863]  
 [864]  
 [865]  
 [866]  
 [867]  
 [868]  
 [869]  
 [870]  
 [871]  
 [872]  
 [873]  
 [874]  
 [875]  
 [876]  
 [877]  
 [878]  
 [879]  
 [880]  
 [881]  
 [882]  
 [883]  
 [884]  
 [885]  
 [886]  
 [887]  
 [888]  
 [889]  
 [890]  
 [891]  
 [892]  
 [893]  
 [894]  
 [895]  
 [896]  
 [897]  
 [898]  
 [899]  
 [900]  
 [901]  
 [902]  
 [903]  
 [904]  
 [905]  
 [906]  
 [907]  
 [908]  
 [909]  
 [910]  
 [911]  
 [912]  
 [913]  
 [914]  
 [915]  
 [916]  
 [917]  
 [918]  
 [919]  
 [920]  
 [921]  
 [922]  
 [923]  
 [924]  
 [925]  
 [926]  
 [927]  
 [928]  
 [929]  
 [930]  
 [931]  
 [932]  
 [933]  
 [934]  
 [935]  
 [936]  
 [937]  
 [938]  
 [939]  
 [940]  
 [941]  
 [942]  
 [943]  
 [944]  
 [945]  
 [946]  
 [947]  
 [948]  
 [949]  
 [950]  
 [951]  
 [952]  
 [953]  
 [954]  
 [955]  
 [956]  
 [957]  
 [958]  
 [959]  
 [960]  
 [961]  
 [962]  
 [963]  
 [964]  
 [965]  
 [966]  
 [967]  
 [968]  
 [969]  
 [970]  
 [971]  
 [972]  
 [973]  
 [974]  
 [975]  
 [976]  
 [977]  
 [978]  
 [979]  
 [980]  
 [981]  
 [982]  
 [983]  
 [984]  
 [985]  
 [986]  
 [987]  
 [988]  
 [989]  
 [990]  
 [991]  
 [992]  
 [993]  
 [994]  
 [995]  
 [996]  
 [997]  
 [998]  
 [999]  
 [1000]

FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 467 467 N-LINKED (GLCNAC. . .)  
 FT DOMAIN 506 611 ASX/GLU-RICH (ACIDIC).  
 FT TRANSMEM 619 637 POTENTIAL.  
 FT MUTAGEN 198 198 S-A: INACTIVATES ENZYME.  
 SQ SEQUENCE 729 AA; 82245 MW; 70583FP279AC02A41 CRC64;  
 Query Match 2.2%; Score 10; DB 1; Length 729;  
 Best Local Similarity 100.0%; Pred. No. 0.092; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 0;  
 QY 434 SEGD DDDDD 443  
 DB 564 SEGD DDDDD 573  
 RESULT 2  
 ID\_NMD2\_YEAST STANDARD: PRT; 1089 AA.  
 AC P38758;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Nonsense-mediated mRNA decay protein 2 (up-frameshift suppressor 2).  
 GN NMD2 OR UPF2 OR IFS1 OR SUI1 OR YHR077C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OC NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95189083; PubMed=7883168;  
 RA He F., Jacobson A.;  
 RT "Identification of a novel component of the nonsense-mediated mRNA  
 decay pathway by use of an interacting protein screen.";  
 Genes Dev. 9:437-454 (1995).  
 [2]  
 [3]  
 [4]  
 [5]  
 [6]  
 [7]  
 [8]  
 [9]  
 [10]  
 [11]  
 [12]  
 [13]  
 [14]  
 [15]  
 [16]  
 [17]  
 [18]  
 [19]  
 [20]  
 [21]  
 [22]  
 [23]  
 [24]  
 [25]  
 [26]  
 [27]  
 [28]  
 [29]  
 [30]  
 [31]  
 [32]  
 [33]  
 [34]  
 [35]  
 [36]  
 [37]  
 [38]  
 [39]  
 [40]  
 [41]  
 [42]  
 [43]  
 [44]  
 [45]  
 [46]  
 [47]  
 [48]  
 [49]  
 [50]  
 [51]  
 [52]  
 [53]  
 [54]  
 [55]  
 [56]  
 [57]  
 [58]  
 [59]  
 [60]  
 [61]  
 [62]  
 [63]  
 [64]  
 [65]  
 [66]  
 [67]  
 [68]  
 [69]  
 [70]  
 [71]  
 [72]  
 [73]  
 [74]  
 [75]  
 [76]  
 [77]  
 [78]  
 [79]  
 [80]  
 [81]  
 [82]  
 [83]  
 [84]  
 [85]  
 [86]  
 [87]  
 [88]  
 [89]  
 [90]  
 [91]  
 [92]  
 [93]  
 [94]  
 [95]  
 [96]  
 [97]  
 [98]  
 [99]  
 [100]  
 [101]  
 [102]  
 [103]  
 [104]  
 [105]  
 [106]  
 [107]  
 [108]  
 [109]  
 [110]  
 [111]  
 [112]  
 [113]  
 [114]  
 [115]  
 [116]  
 [117]  
 [118]  
 [119]  
 [120]  
 [121]  
 [122]  
 [123]  
 [124]  
 [125]  
 [126]  
 [127]  
 [128]  
 [129]  
 [130]  
 [131]  
 [132]  
 [133]  
 [134]  
 [135]  
 [136]  
 [137]  
 [138]  
 [139]  
 [140]  
 [141]  
 [142]  
 [143]  
 [144]  
 [145]  
 [146]  
 [147]  
 [148]  
 [149]  
 [150]  
 [151]  
 [152]  
 [153]  
 [154]  
 [155]  
 [156]  
 [157]  
 [158]  
 [159]  
 [160]  
 [161]  
 [162]  
 [163]  
 [164]  
 [165]  
 [166]  
 [167]  
 [168]  
 [169]  
 [170]  
 [171]  
 [172]  
 [173]  
 [174]  
 [175]  
 [176]  
 [177]  
 [178]  
 [179]  
 [180]  
 [181]  
 [182]  
 [183]  
 [184]  
 [185]  
 [186]  
 [187]  
 [188]  
 [189]  
 [190]  
 [191]  
 [192]  
 [193]  
 [194]  
 [195]  
 [196]  
 [197]  
 [198]  
 [199]  
 [200]  
 [201]  
 [202]  
 [203]  
 [204]  
 [205]  
 [206]  
 [207]  
 [208]  
 [209]  
 [210]  
 [211]  
 [212]  
 [213]  
 [214]  
 [215]  
 [216]  
 [217]  
 [218]  
 [219]  
 [220]  
 [221]  
 [222]  
 [223]  
 [224]  
 [225]  
 [226]  
 [227]  
 [228]  
 [229]  
 [230]  
 [231]  
 [232]  
 [233]  
 [234]  
 [235]  
 [236]  
 [237]  
 [238]  
 [239]  
 [240]  
 [241]  
 [242]  
 [243]  
 [244]  
 [245]  
 [246]  
 [247]  
 [248]  
 [249]  
 [250]  
 [251]  
 [252]  
 [253]  
 [254]  
 [255]  
 [256]  
 [257]  
 [258]  
 [259]  
 [260

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U14974; AAA67724.1; -
DR EMBL: U12137; AAA6521.1; -
DR EMBL: U28158; AAA74948.1; -
DR EMBL: U10556; AAB68893.1; -
DR PIR: S46815; S46815.
DR SCD: S000119; NMD2.
DR InterPro: IPR003890; IF_eIF4G.
DR Pfam: PF02854; MIF4G; 3.
DR SMART: SM00543; MIF4G; 3.
DR DOMAIN 843 975
DR CONFLICT 2 2
DR SEQUENCE 1089 AA; 126746 MW; 138BE725675CBF52 CRC64;
Query Match 2.2%; Score 10; DB 1; Length 1089;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 437 DDDDDDDGEE 446
Db 902 DDDDDDDGEE 911
RESULT 3
MCK1_YEAST
ID MCK1_YEAST STANDARD; PRT; 375 AA.
AC P21965;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein kinase MCK1 (EC 2.7.1.1) (Meiosis and centromere regulatory
DE kinase).
GN MCK1 OR YPK1 OR YNL307C OR N0392.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN 11
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
MEDLINE=91061730; PubMed=1701015;
Dailey D., Schleven G.L., Lim M.-Y., Marguardt H., Gilmore T.,
Thornier J., Martin G.S.;
"Novel yeast protein kinase (YPK1 gene product) is a 40-kilodalton
phosphotyrosyl protein associated with protein-tyrosine kinase
activity."
Mol. Cell. Biol. 10:6244-6256(1990).
(2)
RP SEQUENCE FROM N.A.
RX MEDLINE=91184610; PubMed=2010083;
Najaseborn L., Mitchell A.P.;
"The yeast MCK1 gene encodes a protein kinase homolog that activates
early meiotic gene expression."
Genes Dev. 5:533-548(1991).
(3)
RP SEQUENCE FROM N.A.
RX MEDLINE=91184611; PubMed=2010084;
Shero J.H., Hietter P.;
"A suppressor of a centromere DNA mutation encodes a putative protein
kinase (MCK1)."
Genes Dev. 5:549-560(1991).
(4)
RP SEQUENCE FROM N.A.
RX STRAIN=5288C / FY1676;
MEDLINE=96076632; PubMed=7502583;
Maltani M., Nicoud J.-M., Levesque H., Gallardin C.;

```

```

RT RT "Sequencing analysis of a 24.7 kb fragment of yeast chromosome XIV
RT RT identifies six known genes, a new member of the hexose transporter
RT RT family and ten new open reading frames."
RT RT Yeast 11:1077-1085(1995).
RN (5)
RP SEQUENCE OF 1-16 FROM N.A.
RC STRAIN=5288C / FY1679;
RX MEDLINE=96132033; PubMed=8553702;
RA Maurer K.C.T., Urbanus J.H.M., Planca R.J.;
RT "Sequence analysis of a 30 kb DNA segment from yeast chromosome XIV
RT carrying a ribosomal protein gene cluster, the genes encoding a
RT plasma membrane protein and a subunit of replication factor C, and a
RT novel putative serine/threonine protein kinase gene."
RL Yeast 11:1303-1310(1995).
CC -1- FUNCTION: MAY BE AN AUTOPHOSPHORYLATING TYROSINE KINASE, A
CC BIFUNCTIONAL (SERINE/TYROSINE-SPECIFIC) PROTEIN KINASE, OR A
CC SERINE KINASE THAT IS A SUBSTRATE FOR AN ASSOCIATED TYROSINE
CC KINASE. MCK1 IS A TRANSCRIPTIONAL ACTIVATOR OF IME1, IT STIMULATES
CC SPORE MATURATION, AND PLAY A POSITIVE REGULATORY ROLE IN BOTH
CC MITOTIC CENTROMERE FUNCTION AND ACTIVATION OF EARLY MEIOTIC GENE
CC EXPRESSION.
CC -1- PFM: PHOSPHORYLATED AT TYROSINE AND SERINE IN VIVO.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X55054; CAA38895.1; -
DR EMBL: M55984; AAA34764.1; -
DR EMBL: Z46259; CAA86388.1; -
DR EMBL: Z71583; CAA96236.1; -
DR EMBL: Z71582; CAA96235.1; -
DR PIR: A36354; A36354.
DR PIR: A39622; A39622.
DR PIR: A39623; A39623.
DR HSSP: P24941; IAQ1.
DR SGD: S0005251; MCK1.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00220; STK; 1.
DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
DR TRANSFERASE: Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation.
FT FT SIMILAR 1 8 TO CAMP-DEPENDENT PROTEIN KINASE TYPE 1
FT FT DOMAIN 35 327 (AA 1-8) (IDENTICAL).
FT FT NP_BIND 41 49 PROTEIN KINASE.
FT FT BINDING 68 68 ATP (BY SIMILARITY).
FT FT ACT_SITE 164 164 ATP (BY SIMILARITY).
FT FT MOD_RES 199 199 BY SIMILARITY.
SQ SEQUENCE 375 AA; 43136 MW; BCF0C776B683841B CRC64;
Query Match 2.0%; Score 9; DB 1; Length 375;
Best Local Similarity 100.0%; Pred. No. 0.51;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 145 GVTKICDPRG 153
Db 177 GVTKICDPRG 185
RESULT 4
KIR3_MOUSE
ID KIR3_MOUSE STANDARD; PRT; 502 AA.

```





CC -1- SUBUNIT: TETRAMER OF 2 ALPHA AND 2 BETA CHAINS LINKED BY DISULFIDE  
 CC BONDS. THE ALPHA CHAINS CONTRIBUTE TO THE FORMATION OF THE LIGAND-  
 CC BINDING DOMAIN, WHILE THE BETA CHAINS CARRY THE KINASE DOMAIN (BY  
 CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (by similarity).  
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN  
 CC RECEPTOR SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U72939; AAB17094.1; -.  
 CC HSSP: P06213; IIRK.  
 CC InterPro: IPR000494; EGFR\_L\_domain.  
 CC InterPro: IPR000719; Euk\_kinase.  
 CC InterPro: IPR003961; FN\_III.  
 CC InterPro: IPR002011; Kinaselike.  
 CC InterPro: IPR001245; Tyr\_kinase.  
 CC Pfam: PF00041; fn3; 1.  
 CC Pfam: PF00069; kinase; 1.  
 CC Pfam: PF00757; Furin-like; 1.  
 CC Pfam: PF01030; Recep\_L\_domain; 2.  
 CC PRINTS: PR00109; TYRKINASE.  
 CC ProDom: PD000001; Euk\_kinase; 1.  
 CC SMART: SM00060; FN3; 2.  
 CC SMART: SM00261; TyrKc; 1.  
 CC SMART: SM00219; Tyrc; 1.  
 CC PROSITE: PS00107; PROTEIN KINASE ATP; 1.  
 CC PROSITE: PS00109; PROTEIN KINASE TYR; 1.  
 CC PROSITE: PS00239; RECEPTOR TYR\_KIN\_II; 1.  
 CC PROSITE: PS0011; PROTEIN KINASE\_DOM; 1.  
 CC Transferrase; Tyrosine-protein kinase; Receptor; Transmembrane;  
 CC Glycoprotein; ATP-binding; Phosphorylation; Signal.  
 CC SIGNAL: 1 19  
 CC CHAIN: 20 753  
 CC INSULIN-LIKE RECEPTOR, ALPHA CHAIN  
 CC (POTENTIAL).  
 CC PROPEP 754 757  
 CC CHAIN 758 1390  
 CC INSULIN-LIKE RECEPTOR, BETA CHAIN  
 CC (POTENTIAL).  
 CC DOMAIN 758 966  
 CC TRANSMEM 967 987  
 CC FT 988 1390  
 CC DOMAIN 1025 1301  
 CC NP\_BIND 1031 1039  
 CC BINDING 1059 1059  
 CC ACT\_SITE 1162 1162  
 CC MOD\_RES 1192 1192  
 CC CARBOHYD 196 76  
 CC CARBOHYD 196 196  
 CC CARBOHYD 257 257  
 CC CARBOHYD 322 322  
 CC CARBOHYD 340 340  
 CC CARBOHYD 439 439  
 CC CARBOHYD 457 457  
 CC CARBOHYD 492 492  
 CC CARBOHYD 553 553  
 CC CARBOHYD 765 765  
 CC CARBOHYD 772 772  
 CC CARBOHYD 793 793  
 CC CARBOHYD 796 796  
 CC CARBOHYD 871 871  
 CC CARBOHYD 919 919  
 CC CARBOHYD 933 933  
 CC SEQUENCE 1390 AA; 156831 MW; 6BD2AA382EFP6442 CRC64;

Query Match 2.0%; Score 9; DB 1; Length 1390;  
 Best Local Similarity 100.0%; Pred. No. 1.7;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 191 SYGVVIMEM 199  
 Db 1224 SYGVVIMEM 1232  
 RESULT 8  
 ID UL51\_HCMVA STANDARD; PRT; 157 AA.  
 AC P16792;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 01-FEB-1991 (Rel. 17, Last annotation update)  
 DE Hypothetical protein UL51.  
 GN UL51.  
 OS Human cytomegalovirus (strain AD169).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Betaherpesvirinae; Cytomegalovirus.  
 OX NCBI\_TaxID=10360;  
 RN [1]  
 RP MEDLINE=90269039; PubMed=2161319;  
 RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,  
 RA Horsnell T., Hutchinson C.A. III, Kouzarides T., Martignetti J.A.,  
 RA Preddie B., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.,  
 RT "Analysis of the protein-coding content of the sequence of human  
 RT cytomegalovirus strain AD169."  
 RT Curr. Top. Microbiol. Immunol. 154:125-169 (1990).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X17403; CA35410.1; -.  
 CC PIR: S09814; S09814.  
 CC InterPro: IPR005208; Herpes\_UL33.  
 CC Pfam: PF03581; Herpes\_UL33; 1.  
 CC KW Hypothetical protein.  
 CC SEQUENCE 157 AA; 16968 MW; 5999035AC654517D CRC64;  
 SO  
 Query Match 1.8%; Score 8; DB 1; Length 157;  
 Best Local Similarity 100.0%; Pred. No. 2.3;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 440 DDDGDEE 447  
 Db 12 DDDGDEE 19  
 RESULT 9  
 ID 2SS\_SOYBN STANDARD; PRT; 158 AA.  
 AC P13594;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 2S albumin precursor (Gm2S-1) [Contains: Aspartic acid-rich peptide;  
 DE 8 kDa methionine-rich protein (8 kDa MRP)].  
 OS Glycine max (Soybean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 OX NCBI\_TaxID=3847;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Hodgson 78; TISSUE=Cotyledon;  
 RA Galvez A.F., Revilla M.J.R., de Lumen B.O.;  
 RA "A novel methionine-rich protein from soybean cotyledon: cloning and



RT Characterization of cDNA."  
 RL (1n) Plant Gene Register PGR97-103.  
 RN  
 RP SEQUENCE OF 22-64.  
 RC TISSUE=Seed;  
 RX MEDLINE=87280104; PubMed=3611081;  
 RA Odani S., Koide T., Ono T.;  
 RT "Antio acid sequence of a soybean (glycine max) seed polypeptide  
 RL having a poly(L-aspartic acid) structure."  
 RL J. Biol. Chem. 262:10502-10505 (1987).  
 RN  
 RP SEQUENCE OF 82-96.  
 RA Revilla M.J., Galvez A.F., Krenz D.C., de Lumen B.O.;  
 RT "An 8 kDa methionine-rich protein (MRP) from soybean (glycine max)  
 RL coryledon: identification, purification and N-terminal sequence."  
 CC J. Agric. Food Chem. 44:2930-2935 (1996).  
 CC  
 CC -1- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.  
 CC -1- SUBUNIT: THE PROTEIN CONSISTS OF TWO CHAINS LINKED BY 2 DISULFIDE  
 CC BONDS.  
 CC  
 CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AF005030; AAB71140.1; -  
 DR PIR: A28485; A28485.  
 DR InterPro: IPR003612; AAI.  
 DR InterPro: IPR001768; Try/amy1\_inhbr.  
 DR Pfam: PF00234; try\_alpha\_amy1; 1.  
 DR SMART: SM00499; AAI; 1.  
 KW Signal; Seed storage protein.  
 FT SIGNAL 1 21  
 FT CHAIN 22 64 2S ALBUMIN, SMALL CHAIN (ASPARTIC ACID-  
 FT RICH PEPTIDE).  
 FT PROPEP 65 81  
 FT CHAIN 82 158 2S ALBUMIN, LARGE CHAIN (8 KDA MRP).  
 FT SITE 54 56 CELL ATTACHMENT SITE (POTENTIAL).  
 FT DOMAIN 56 64 POLY-ASP.  
 FT DOMAIN 88 91 POLY-GLU.  
 SQ SEQUENCE 158 AA; 18460 MW; C1F42723B0F62D81 CRC64;  
 Query Match 1.8%; Score 8; DB 1; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 2.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 436 GDDDDDD 443  
 DB 55 GDDDDDD 62  
 RESULT 10  
 REGO\_BP82 STANDARD; PRT; 229 AA.  
 AC P13870;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Antitermination protein Q.  
 GN O.  
 OS Bacteriophage 82.  
 CC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;  
 OC Lambda-like viruses.  
 OX NCBI\_TaxID=10705;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87308148; PubMed=3624233;  
 RA Goliger J.A., Roberts J.W.;  
 RT "Bacteriophage 82 gene Q and Q protein. Sequence, overproduction, and

RT activity as a transcription antiterminator in vitro."  
 RL J. Biol. Chem. 262:11721-11725 (1987).  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96196428; PubMed=8648624;  
 RA Mahdi A.A., Sharples G.J., Mandal T.N., Lloyd R.G.;  
 RT "Holiday junction resolvases encoded by homologous ruva genes in  
 RT Escherichia coli K-12 and phage 82."  
 RL J. Mol. Biol. 257:561-573 (1996).  
 CC  
 CC -1- FUNCTION: POSITIVELY REGULATE EXPRESSION OF THE PHAGE LATE GENE  
 CC OPERONS. BACTERIAL HOST RNA POLYMERASE MODIFIED BY ANTITERMINATION  
 CC PROTEINS TRANSCRIBES THROUGH TERMINATION SITES THAT OTHERWISE  
 CC PREVENT EXPRESSION OF THE REGULATED GENES.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: J02803; AAA32298.1; -  
 DR PIR: A29791; POB82.  
 DR PIR: A29791; POB82.  
 KW Transcription regulation; Transcription termination; DNA-binding.  
 FT DNA BIND 14 32 POTENTIAL.  
 SQ SEQUENCE 229 AA; 26400 MW; E6BDB8E81498667 CRC64;  
 Query Match 1.8%; Score 8; DB 1; Length 229;  
 Best Local Similarity 100.0%; Pred. No. 3.3;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 290 TIERLKL 297  
 DB 142 TIERLKL 149  
 RESULT 11  
 MCML\_YEAST STANDARD; PRT; 286 AA.  
 AC P11746;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Phomone receptor transcription factor (GRW/PRTF protein).  
 GN MCML OR FUN80 OR YMR043W OR YW532.08.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90249735; PubMed=2159934;  
 RA Ammerer G.;  
 RT "Identification, purification, and cloning of a polypeptide  
 RT (PRTF/GRW) that binds to mating-specific promoter elements in  
 RT yeast."  
 RL Genes Dev. 4:299-312 (1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89141759; PubMed=3066908;  
 RA Paasmore S., Maine G.T., Elble R., Christ C., Ty B.K.;  
 RT "Saccharomyces cerevisiae protein involved in plasmid maintenance is  
 RT necessary for mating of MAT alpha cells."  
 RL J. Mol. Biol. 204:593-606 (1988).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88030692; PubMed=3311883;  
 RA Dubois E., Bercy J., Descamps F.;  
 RT "Characterization of two new genes essential for vegetative growth in  
 RT Saccharomyces cerevisiae: nucleotide sequence determination and  
 RT chromosome mapping.";

Gene 55:265-275(1987).

[4]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=5288c / AB972;  
 RA Odeil C., Bowman S., Barrell B.G., Rajandream M.A.;  
 RL Submitted (Feb-1995) to the EMBL/GenBank/DBJ databases.

[5]  
 RN X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF COMPLEX WITH MAT2.  
 RP MEDLINE=98140702; PubMed=9490409;  
 RA Tan S., Richmond T.J.;  
 RL "Crystal structure of the yeast Matalpha2/Mcm1/DNA ternary complex.",  
 Nature 391:660-666(1998).

CC -1- FUNCTION: INTERACTS WITH THE ALPHA-2 REPRESSOR OR WITH THE  
 ALPHA-1 ACTIVATOR THEREBY REGULATING THE EXPRESSION OF MATING-  
 TYPE-SPECIFIC GENES.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION  
 FACTORS.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: X52453; CAA36691.1; -  
 DR EMBL: X14187; CAA32389.1; -  
 DR EMBL: X17511; AAA34609.1; -  
 DR EMBL: Z48502; CAA88409.1; -  
 DR PIR: A34599; A34599.  
 DR PDB: 1MMW; 18-MAR-98.  
 DR TRANSFAC: T00500; -  
 DR TRANSFAC: T00501; -  
 DR SGD: S0004646; MCM1.  
 DR InterPro: IPR002100; TF\_MADSbox.  
 DR Pfam: PF00319; SRF-TF.1.  
 DR PRINTS: PRO0404; MADSDOMAIN.  
 DR SMART: SM00432; MADS.1.  
 DR PROSITE: PS00350; MADS\_BOX\_1; 1.  
 DR PROSITE: PS00066; MADS\_BOX\_2; 1.  
 DR Transcription regulation; DNA-binding; Nuclear protein; 3D-structure.

CC KW DOMAIN 18 72 MADS.  
 FT DOMAIN 98 120 ASP/GLU-RICH (ACIDIC).  
 FT CONFLICT 9 9 P -> S (IN REF. 3).  
 FT CONFLICT 37 37 S -> F (IN REF. 3).  
 FT CONFLICT 156 157 GA -> AR (IN REF. 3).  
 FT CONFLICT 158 286 MISSING (IN REF. 3).  
 FT SEQUENCE 286 AA; 32802 MW; PD75B4B549E9E3B CRC64;

CC Query Match 1.8%; Score 8; DB 1; Length 286;  
 CC Best Local Similarity 100.0%; Pred. No. 4;  
 CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 436 GDDDDDD 443  
 Db 108 GDDDDDD 115

RESULT 12  
 OSTP\_MOUSE STANDARD; PRT; 294 AA.  
 ID OSTP\_MOUSE  
 AC P10923; P19008; (Rel. 11, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 01-JUL-1989 (Rel. 11, Last annotation update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Osteopontin precursor (Bone sialoprotein 1) (Minopontin) (Early T  
 DE lymphocyte activation 1 protein) (Secreted phosphoprotein 1) (SPP-1)  
 DE (2AR) (Calcium oxalate crystal growth inhibitor protein).  
 DE SPP1 OR SPP-1 OR OP OR ETA-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Liver;  
 RA MEDLINE=9035443; PubMed=2387863;  
 RA Miyazaki Y., Setoguchi M., Yoshida S.Y., Akizuki S., Yamamoto S.;  
 RT "The mouse osteopontin gene. Expression in monocytic lineages and  
 RT complete nucleotide sequence.",  
 RL J. Biol. Chem. 265:14432-14438(1990).

[2]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Macrophage;  
 RC MEDLINE=89263742; PubMed=2726465;  
 RA Miyazaki Y., Setoguchi M., Yoshida S., Higuchi Y., Akizuki S.,  
 RA Yamamoto S.;  
 RT "Nucleotide sequence of cDNA for mouse osteopontin-like protein.",  
 RL Nucleic Acids Res. 17:3298-3298(1989).

[3]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=89255479; PubMed=2722855;  
 RC Craig A.M., Smith J.H., Denhardt D.T.;  
 RA "Osteopontin, a transformation-associated cell adhesion  
 RT phosphoprotein, is induced by 12-O-tetradecanoylphorbol 13-acetate in  
 RT mouse epidermis.",  
 RL J. Biol. Chem. 264:9682-9689(1989).

[4]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=89310352; PubMed=2787378;  
 RA Patarca R., Freeman G.J., Sinch R.P., Wei F.-Y., Durfee T.,  
 RA Blatner F., Regnier D.C., Kozak C.A., Mock B.A., Morse H.C. III,  
 RA Jerrils T.R., Cantor H.;  
 RT "Structural and functional studies of the early T lymphocyte  
 RT activation 1 (Eta-1) gene. Definition of a novel T cell-dependent  
 RT response associated with genetic resistance to bacterial infection.",  
 RL J. Exp. Med. 170:145-161(1989).

[5]  
 RN SEQUENCE OF 17-37.  
 RP TISSUE=Kidney;  
 RC MEDLINE=93034441; PubMed=1414495;  
 RA Worcester E.M., Blumenthal S.S., Beshensky A.M., Lewand D.L.;  
 RT "The calcium oxalate crystal growth inhibitor protein produced by  
 RT mouse kidney cortical cells in culture is osteopontin.",  
 RL J. Bone Miner. Res. 7:1029-1036(1992).

CC -1- FUNCTION: BINDS TIGHTLY TO HYDROXYAPATITE. APPEARS TO FORM AN  
 CC INTEGRAL PART OF THE MINERALIZED MATRIX. PROBABLY IMPORTANT TO  
 CC CELL-MATRIX INTERACTION.  
 CC -1- SUBUNIT: LIGAND FOR INTEGRIN ALPHA-V/BETA-3.  
 CC -1- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES.  
 CC -1- PTM: N- AND O-GLYCOSYLATED.  
 CC -1- SIMILARITY: BELONGS TO THE OSTEOPONTIN FAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: X51834; CAA36132.1; -  
 DR EMBL: X13986; CAA32165.1; -  
 DR EMBL: J04806; AAA57265.1; -  
 DR EMBL: X16151; CAA34276.1; -  
 DR PIR: J01015; J01015.  
 DR PIR: A33853; A33853.  
 DR PIR: A37818; A37818.  
 DR PIR: S12064; S12064.  
 DR MGI: MGI:98389; Spp1.  
 DR InterPro: IPR002038; Osteopontin.  
 DR Pfam: PF00865; Osteopontin; 1.  
 DR PRINTS: PRO0216; OSTEOPONTIN.  
 DR SMART: SM00017; OSTEO; 1.

DR PROSITE; PS00884; OSTEOPONTIN; 1.  
 KW Glycoprotein; Sialic acid; Bone; Cell adhesion; Phosphorylation;  
 FT SIGNAL 1 16  
 FT CHAIN 17 294 POTENTIAL.  
 FT SITE 144 146 OSTEOPONTIN.  
 FT CARBOHYD 78 78 CELL ATTACHMENT SITE.  
 FT CONFLICT 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 99 99 L -> P (IN REF. 4).  
 FT CONFLICT 122 122 E -> G (IN REF. 3).  
 FT CONFLICT 122 122 V -> F (IN REF. 2).  
 SQ SEQUENCE 294 AA; 32459 MW; 9DSF32D67AB53EA CRC64;  
 Query Match  
 Best Local Similarity 1.8%; Score 8; DB 1; Length 294;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 437 DDDDDDDG 444  
 Db 88 DDDDDDDG 95  
 RESULT 13  
 ID OSTP\_RAT STANDARD; PRT; 317 AA.  
 AC P08721;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Osteopontin precursor (Bone sialoprotein 1) (Secreted phosphoprotein 1) (SPP-1).  
 GN SPP1 OR SPP-1 OR 2B7.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=93054745; PubMed=1429723;  
 RA Singh K., Mukherjee A.B., De Vouge M.W., Mukherjee B.B.;  
 RT "Differential processing of osteopontin transcripts in rat kidney-  
 and osteoblast-derived cell lines.";  
 RL J. Biol. Chem. 267:23847-23851(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87067405; PubMed=3024151;  
 RA Oldberg A., Franzen A., Heinegaard D.;  
 RT "Cloning and sequence analysis of rat bone sialoprotein (osteopontin)  
 cDNA reveals an Arg-Gly-Asp cell-binding sequence.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:8819-8823(1986).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Smooth muscle;  
 RX MEDLINE=94013467; PubMed=8406622;  
 RA Giachelli C.M., Bae N., Almeida M., Denhardt D.T., Alpers C.E.,  
 RA Schwartz S.M.;  
 RT "Osteopontin is elevated during neointima formation in rat arteries  
 and is a novel component of human atherosclerotic plaques.";  
 RL J. Clin. Invest. 92:1686-1696(1993).  
 CC -1- FUNCTION: BINDS TIGHTLY TO HYDROXYAPATITE. APPEARS TO FORM AN  
 INTEGRAL PART OF THE MINERALIZED MATRIX. PROBABLY IMPORTANT TO  
 CELL-MATRIX INTERACTION.  
 CC -1- SUBUNIT: LIGAND FOR INTEGRIN ALPHA-V/BETA-3.  
 CC -1- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES.  
 CC -1- SIMILARITY: BELONGS TO THE OSTEOPONTIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----  
 DR EMBL; M9252; AAA41765.1; -  
 DR EMBL; M4656; AAA41762.1; -  
 DR PIR; A25917; A25917.  
 DR InterPro; IPR002038; Osteopontin.  
 DR Pfam; PF00865; Osteopontin; 1.  
 DR PRINTS; PR00216; Osteopontin.  
 DR SMART; SM00017; OSTEO; 1.  
 DR PROSITE; PS00884; OSTEOPONTIN; 1.  
 KW Glycoprotein; Sialic acid; Bone; Cell adhesion; Phosphorylation;  
 FT SIGNAL 1 22  
 FT CHAIN 23 317 POTENTIAL.  
 FT DOMAIN 86 96 OSTEOPONTIN.  
 FT SITE 144 146 POLY-ASP.  
 FT CARBOHYD 79 79 CELL ATTACHMENT SITE.  
 FT CONFLICT 8 8 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 8 8 F -> L (IN REF. 2).  
 SQ SEQUENCE 317 AA; 34963 MW; 73CB5C21FF62310 CRC64;  
 Query Match  
 Best Local Similarity 1.8%; Score 8; DB 1; Length 317;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 437 DDDDDDDG 444  
 Db 88 DDDDDDDG 95  
 RESULT 14  
 ID RLAO\_HALCU STANDARD; PRT; 352 AA.  
 AC P1706;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Acidic ribosomal protein P0 homolog (L10E).  
 GN RPLP0.  
 OS Halobacterium cutribdum.  
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
 OC Halobacteriaceae; Halobacterium.  
 OX NCBI\_TaxID=2242;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=34001;  
 RX MEDLINE=89305527; PubMed=2743981;  
 RA Shimmin L.C., Dennis P.P.;  
 RT "Characterization of the L11, L1, L10 and L12 equivalent ribosomal  
 RT protein gene cluster of the halophilic archaeobacterium Halobacterium  
 cutribdum.";  
 RL EMBO J. 8:1225-1235(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 CC -1- FUNCTION: RIBOSOMAL PROTEIN P0 IS THE FUNCTIONAL EQUIVALENT  
 CC OF E-COLI PROTEIN L10.  
 CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----  
 DR EMBL; X15078; CAA33180.1; -  
 DR PIR; S04120; RSHSOC.  
 DR InterPro; IPR001790; Ribosomal L10.  
 DR Pfam; PF00466; Ribosomal\_L10; 1.  
 DR RiboSomal protein.  
 SQ SEQUENCE 352 AA; 37199 MW; 2AE662E79A9BD02D CRC64;  
 Query Match  
 Best Local Similarity 1.8%; Score 8; DB 1; Length 352;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



RA Gileadi O., Kornberg R.D.,  
 RT "yeast TFIIIE. Cloning, expression, and homology to vertebrate  
 RT proteins.";  
 RL J. Biol. Chem. 269:27549-27553(1994).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RA Rieger M.,  
 RL Submitted (MAR-1994) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: RECRUITS TFIIH TO THE INITIATION COMPLEX AND STIMULATES  
 CC THE RNA POLYMERASE II C-TERMINAL DOMAIN KINASE AND DNA-DEPENDENT  
 CC ATPASE ACTIVITIES OF TFIIH. BOTH TFIIH AND TFIIIE ARE REQUIRED FOR  
 CC PROMOTER CLEARANCE BY RNA POLYMERASE (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE TFIIIE ALPHA SUBUNIT FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U12825; AAA62665.1; -;  
 CC EMBL: Z28028; CAA81863.1; -;  
 CC PIR: S37845; S37845; -;  
 CC TRASNSEC: T02239; -;  
 CC SGD: S0001511; TPA1.  
 CC InterPro: IPR002853; TFIIIE\_alpha.  
 CC Pfam: PF02002; TFIIIE\_alpha.  
 CC SMART: SM00531; TFIIIE\_1.  
 CC K2M Transcription regulation; Nuclear protein; Zinc-finger.  
 CC ZN\_FING 124 152 C4-TYPE (POTENTIAL).  
 CC FT DOMAIN 374 482 ASF/GLU-RICH (HIGHLY ACIDIC).  
 CC FT DOMAIN 374 392 POLY-GLU.  
 CC SQ SEQUENCE 482 AA; 54742 MW; 3E789DFC4247EF8A CRC64;  
 CC  
 CC Query Match 1.8%; Score 8; DB 1; Length 482;  
 CC Best Local Similarity 100.0%; Pred. No. 6.5;  
 CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 436 GDDDDDD 443  
 CC DB 464 GDDDDDD 471  
 CC  
 CC RESULT 18  
 CC K4 HUMAN  
 CC STK4 HUMAN STANDARD; PRT; 487 AA.  
 CC NC 013043 \* O15802; Q9NT24;  
 CC DT 16-OCT-2001 (Rel. 40; Created)  
 CC DT 16-OCT-2001 (Rel. 40; Last sequence update)  
 CC DE 16-OCT-2001 (Rel. 40; Last annotation update)  
 CC DE Serine/threonine protein kinase 4 (EC 2.7.1.37) (STE20-like kinase  
 CC MST1) (MST-1) (Mammalian STE20-like protein kinase 1)  
 CC DE (Serine/threonine protein kinase Krs-2).  
 CC GN STK4 OR MST1.  
 CC OS Homo sapiens (Human).  
 CC CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 CC OK NCBI\_TaxID=9606;  
 CC [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RA MEDLINE=95394928; PubMed=765586;  
 CC RT "Cloning and characterization of a human protein kinase with homology  
 CC RT to Ste20.";  
 CC RL J. Biol. Chem. 270:21695-21700(1995).  
 CC RN (2)  
 CC RP SEQUENCE FROM N.A.  
 CC RA MEDLINE=96413604; PubMed=8816758;  
 CC RA Taylor L.K., Wang H.C., Erikson R.L.;  
 CC RT "Newly identified stress-responsive protein kinases, Krs-1 and Krs-

RT 2.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:10099-10104(1996).  
 RN (3)  
 RP SEQUENCE OF 1-435 FROM N.A.  
 RA Laird G.;  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -1- ENZYME REGULATION: THE C-TERMINAL NON-CATALYTIC REGION INHIBITS  
 CC THE KINASE ACTIVITY.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED.  
 CC -1- PTM: AUTOPHOSPHORYLATED ON SERINE AND THREONINE RESIDUES.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC STE20 SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U12827; AAA83254.1; -;  
 CC EMBL: U60207; AAB17262.1; -;  
 CC EMBL: AL109839; CAB89421.1; -;  
 CC HSP: P24941; IHCL.  
 CC Genew: HGNC:11408; STK4.  
 CC MIM: 604965; -;  
 CC InterPro: IPR000719; Btk\_pkinase.  
 CC InterPro: IPR002230; Ser\_thr\_pkinase.  
 CC Pfam: PF00069; Pkinase; 1.  
 CC PRINTS: PR00109; TYRKINASE.  
 CC PRODOM: PD000001; Btk\_pkinase; 1.  
 CC SMART: SM00220; S\_TKC\_1.  
 CC DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.  
 CC DR PROSITE: PS00108; PROTEIN KINASE-ST; FALSE NEG.  
 CC DR PROSITE: PS50011; PROTEIN KINASE-DOM; 1.  
 CC KM Transferrase; Serine/threonine-protein kinase; ATP-binding.  
 CC FT DOMAIN 30 281  
 CC FT NP\_BIND 36 44  
 CC FT BINDING 59 59  
 CC FT ACT\_SITE 149 149  
 CC FT DOMAIN 373 378  
 CC FT CONFLICT 222 222  
 CC FT CONFLICT 312 312  
 CC FT CONFLICT 312 312  
 CC SQ SEQUENCE 487 AA; 55630 MW; 150758BEC5F77D5C CRC64;  
 CC  
 CC Query Match 1.8%; Score 8; DB 1; Length 487;  
 CC Best Local Similarity 100.0%; Pred. No. 6.6;  
 CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 172 WMAPVVIQ 179  
 CC DB 190 WMAPVVIQ 197  
 CC  
 CC RESULT 19  
 CC STK3 HUMAN  
 CC ID STK3 HUMAN STANDARD; PRT; 491 AA.  
 CC AC 013168; O15801; Q15445;  
 CC DT 16-OCT-2001 (Rel. 40; Created)  
 CC DT 16-OCT-2001 (Rel. 40; Last sequence update)  
 CC DE 16-OCT-2001 (Rel. 40; Last annotation update)  
 CC DE Serine/threonine protein kinase 3 (EC 2.7.1.37) (STE20-like kinase  
 CC MST2) (MST-2) (Mammalian STE20-like protein kinase 2)  
 CC DE (Serine/threonine protein kinase Krs-1).  
 CC GN STK3 OR MST2.  
 CC OS Homo sapiens (Human).  
 CC CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 CC OK NCBI\_TaxID=9606;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96144292; PubMed=8566796;  
 RA Creasy C.L., Chernoff U.;  
 RT "Cloning and characterization of a member of the MST subfamily of  
 RL Ste20-like kinases.";  
 RL Gene 167:303-306 (1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96413604; PubMed=8816758;  
 RA Taylor L.K., Wang H.C., Erikson R.L.;  
 RT "Newly identified stress-responsive protein kinases, Krs-1 and Krs-  
 RL 2.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:10099-10104 (1996).  
 RN [3]  
 RP SEQUENCE OF 96-203 FROM N.A.  
 RX MEDLINE=94100173; PubMed=8274451;  
 RA Schultz S.J., Nigg B.A.;  
 RT "Identification of 21 novel human protein kinases, including 3 members  
 RL of a family related to the cell cycle regulator hima of Aspergillus  
 RL nidulans.";  
 RL Cell Growth Differ. 4:821-830 (1993).  
 CC -1- FUNCTION: OXIDANT STRESS-ACTIVATED SERINE/THREONINE KINASE THAT  
 CC MAY PLAY A ROLE IN THE RESPONSE TO ENVIRONMENTAL STRESS (BY  
 CC SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN ADULT KIDNEY,  
 CC SKELETAL AND PLACENTA TISSUES AND AT VERY LOW LEVELS IN ADULT  
 CC HEART, LUNG AND BRAIN TISSUES.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC STE20 SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: U26424; AAC50386.1; -  
 CC EMBL: U60206; AAB17261.1; -  
 CC EMBL: Z25422; CAA80909.1; -  
 CC HSP: P24941; 1HCL.  
 CC Genew: HGNC:11406; STK3.  
 CC MIM: 605030; -  
 CC InterPro: IPR000719; Euk\_pkinase.  
 CC InterPro: IPR002290; Ser\_thr\_pkinase.  
 CC Pfam: PF00069; pkinase; 1.  
 CC ProDom: PD000001; Euk\_pkinase; 1.  
 CC SMART: SM00220; S\_TKc; 1.  
 CC PROSITE: PS00107; PROTEIN KINASE ATP; 1.  
 CC PROSITE: PS00108; PROTEIN KINASE ST; FALSE\_NEG.  
 CC PROSITE: PS50011; PROTEIN KINASE DOM; 1.  
 CC Transferrase; Serine/threonine-protein kinase; ATP-binding.  
 CC KW DOMAIN 27 278 PROTEIN KINASE.  
 CC FT NP BIND 33 41 ATP (BY SIMILARITY).  
 CC FT BINDING 56 56 ATP (BY SIMILARITY).  
 CC FT ACT SITE 146 146 BY SIMILARITY.  
 CC FT DOMAIN 308 375 POLY-GLU.  
 CC FT CONFLICT 96 98 WIV -> YLY (IN REF. 3).  
 CC FT CONFLICT 121 121 D -> Y (IN REF. 3).  
 CC FT CONFLICT 203 203 D -> G (IN REF. 3).  
 CC FT CONFLICT 303 303 D -> E (IN REF. 2).  
 CC FT CONFLICT 332 334 GEC -> ESV (IN REF. 2).  
 CC SEQUENCE 491 AA; 56261 MW; 9CA3B0644F3C14A9 CRC64;  
 SQ  
 Query Match 1.8%; Score 8; DB 1; Length 491;  
 Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 172 WMAPEVIO 179  
 Db 187 WMAPEVIO 194  
 RESULT 20  
 MK7 MOUSE STANDARD; PRT; 579 AA.  
 AC 062073;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Mitogen-activated protein kinase kinase 7 (EC 2.7.1.-)  
 DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-  
 DE activated kinase 1).  
 GN MAPK7 OR TAK1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCB1\_TaxID=10090;  
 RX SEQUENCE FROM N.A.  
 RX MEDLINE=96123277; PubMed=8533096;  
 RA Yamaguchi K., Shirakabe K., Shibuya H., Irie K., Ohishi I., Ueno N.,  
 RA Taniguchi T., Nishida E., Matsumoto K.;  
 RT "Identification of a member of the MAPKK family as a potential  
 RL mediator of TGF-beta signal transduction.";  
 RL Science 270:2008-2011 (1995).  
 CC -1- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKKS.  
 CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B  
 CC ACTIVATION.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MAP KINASE SUBFAMILY.  
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C210RF7.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: D76446; BAA11184.1; -  
 CC HSP: P08631; IAD5.  
 CC MGD: MGI:1346877; Map3k7.  
 CC InterPro: IPR000719; Euk\_pkinase.  
 CC InterPro: IPR004040; STY\_pkinase.  
 CC InterPro: IPR002290; Ser\_thr\_pkinase.  
 CC Pfam: PF00069; pkinase; 1.  
 CC ProDom: PD000001; Euk\_pkinase; 1.  
 CC SMART: SM00221; STYKc; 1.  
 CC PROSITE: PS00107; PROTEIN KINASE ATP; 1.  
 CC PROSITE: PS00108; PROTEIN KINASE ST; 1.  
 CC PROSITE: PS50011; PROTEIN KINASE DOM; 1.  
 CC Transferrase; Serine/threonine-protein kinase; ATP-binding.  
 CC KW DOMAIN 8 16 PROTEIN KINASE.  
 CC FT NP BIND 36 291 ATP (BY SIMILARITY).  
 CC FT BINDING 42 50 ATP (BY SIMILARITY).  
 CC FT ACT SITE 156 156 BY SIMILARITY.  
 CC SEQUENCE 579 AA; 64227 MW; 97C8F6F3C8E283EE CRC64;  
 SQ  
 Query Match 1.8%; Score 8; DB 1; Length 579;  
 Best Local Similarity 100.0%; Pred. No. 7.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 146 VLKICDPG 153  
 Db 170 VLKICDPG 177  
 RESULT 21

YMW2\_YEAST STANDARD; PRT; 590 AA.  
 ID YMW2\_YEAST  
 AC 005021;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Hypothetical 67.6 kDa protein in MRPL44-MTF1 intergenic region.  
 GN YMR227C OR YMR959.09C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 NC NCB1\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SG288C / AB972;  
 RL Skelton U., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;  
 CC Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; Z49939; CAA90198.1; -  
 DR TRANSFAC; T03088;  
 DR SGD; S0004840; YMR227C.  
 KW Hypothetical protein; Coiled coil.  
 FT DOMAIN 50 53 POLY-LYS.  
 FT DOMAIN 200 203 POLY-GLU.  
 FT DOMAIN 368 373 POLY-GLU.  
 FT DOMAIN 413 421 POLY-ASP.  
 FT DOMAIN 517 531 POLY-GLU.  
 FT DOMAIN 427 549 COILED COIL (POTENTIAL).  
 SO SEQUENCE 590 AA; 67555 MW; C014E7419B0B1C61 CRC64;  
 Query Match 1.8%; Score 8; DB 1; Length 590;  
 Best Local Similarity 100.0%; Pred. No. 7.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 437 DDDDDDDG 444  
 DB 413 DDDDDDDG 420  
 RESULT 22  
 M3K7\_HUMAN STANDARD; PRT; 606 AA.  
 AC 043318; O43317; O43319;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Mitogen-activated protein kinase kinase 7 (EC 2.7.1.-)  
 DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-  
 DE activated kinase 1).  
 GN MAP3K7 OR TAK1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NC NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.  
 RC TISSUE=lung;  
 RX MEDLINE=86135801; PubMed=9480845;  
 RA Sakurai H., Shigemori N., Hasegawa K., Sugita T.;  
 RT "TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an  
 RT NF-kappa B-inducing kinase-independent mechanism."  
 RL Biochem. Biophys. Res. Commun. 243:545-549(1998).  
 CC -1- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKs.  
 CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B  
 CC ACTIVATION.

CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: 1A, 1B (SHOWN HERE) AND 1C; ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MAP KINASE KINASE KINASE SUBFAMILY.  
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C210RF7.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AB009357; BA25026.1; -  
 DR EMBL; AB009356; BA25025.1; -  
 DR EMBL; AB009358; BA25027.2; -  
 DR HSSP; P06631; 1AD5  
 DR Genew; HGNC:6859; MAP3K7.  
 DR MIM; 602614; -  
 DR InterPro; IPR000719; Btk\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Btk\_pkinase; 1.  
 DR SMART; SM00221; STYKc; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;  
 FT DOMAIN 8 14 POLY-SER.  
 FT DOMAIN 36 221 PROTEIN KINASE.  
 FT NP BIND 42 50 ATP (BY SIMILARITY).  
 FT BINDING 63 63 ATP (BY SIMILARITY).  
 FT ACT SITE 156 156 BY SIMILARITY.  
 FT VARSPLIC 404 430 MISSING (IN ISOFORM 1A).  
 FT VARSPLIC 509 518 MISSING (IN ISOFORM 1C).  
 FT VARSPLIC 519 606 MISSING (IN ISOFORM 1C).  
 SO SEQUENCE 606 AA; 67196 MW; 3DBF8147CD174013 CRC64;  
 Query Match 1.8%; Score 8; DB 1; Length 606;  
 Best Local Similarity 100.0%; Pred. No. 8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 146 VLKICDPG 153  
 DB 170 VLKICDPG 177  
 RESULT 23  
 E2H2\_HUMAN STANDARD; PRT; 746 AA.  
 AC 015910; O92857; Q15755;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Enhancer of zeste homolog 2 (ENX-1).  
 GN E2H2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NC NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=97124843; PubMed=8954776;  
 RA Chen H., Roessler C., Antonarakis S.E.;  
 RT "Cloning of a human homolog of the Drosophila enhancer of zeste gene  
 RT (E2H2) that maps to chromosome 21q22.2."  
 RL Genomics 38:30-37(1996).  
 CC [2]  
 CC SEQUENCE FROM N.A.

RX MEDLINE=97357309; PubMed=9224638;  
 RA Laible G., Wolf A., Dorn R., Reuter G., Nislow C., Lebersorger A.,  
 RA Popkin D., Pillus L., Jenwein T.;  
 RT "Mammalian homologues of the Polycomb-group gene Enhancer of zeste  
 RT mediate gene silencing in Drosophila heterochromatin and at S.  
 RT cerevisiae telomeres.";  
 RL EMO. J. 16:3219-3232 (1997).  
 [3]  
 RN SEQUENCE OF 134-746 FROM N.A.  
 RP MEDLINE=96220494; PubMed=8649418;  
 RX Hobert O., Jallil B., Ullrich A.;  
 RA "Interaction of Vav with ENX-1, a putative transcriptional regulator  
 RT of homeobox gene expression".  
 RT Mol. Cell. Biol. 16:3066-3073 (1996).  
 [4]  
 RN ATRX BINDING.  
 RP MEDLINE=98167853; PubMed=9499421;  
 RX Cardoso C., Timst S., Villard L., Khrestchatsky M., Fontes M.,  
 RA Colleaux L.;  
 "Specific interaction between the XNP/ATRX-X gene product and the SET  
 domain of the human EZH2 protein.";  
 Hum. Mol. Genet. 7:679-684 (1998).  
 CC -1- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF GENE TRANSCRIPTION  
 AND CHROMATIN STRUCTURE.  
 CC -1- SUBUNIT: BINDS ATRX VIA THE SET DOMAIN (PROBABLE).  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY TISSUES.  
 CC -1- SIMILARITY: CONTAINS 1 SET DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE EZ FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X96653; CAA64955.1; -;  
 DR EMBL: U61145; AAC51520.1; -;  
 DR EMBL: U52965; AAC50591.1; -;  
 DR TRANSFAC: T04888; -;  
 DR Genew; HGNC:3527; EZH2.  
 DR MIM: 601573; -;  
 DR InterPro: IPR001005; Myb\_DNA\_binding.  
 DR InterPro: IPR001214; SET.  
 DR Pfam; PF00856; SET; 1.  
 DR SMART; SM00395; SANT; 2.  
 DR SMART; SM00317; SET; 1.  
 DR PROSITE; PS50280; SET; 1.  
 KW Transcription regulation; Nuclear protein; DNA-binding.  
 FT DOMAIN 490 495 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT DOMAIN 523 605 CYS-RICH.  
 FT DOMAIN 618 729 SET.  
 FT DOMAIN 724 724 F -> L (IN REF. 1).  
 FT CONFLICT 224 224 F -> V (IN REF. 1).  
 FT CONFLICT 724 724 F -> V (IN REF. 1).  
 SQ SEQUENCE 746 AA; 85363 MW; 1B5029BB9D509B55 CRC64;  
 Query Match 1.8%; Score 8; DB 1; Length 746;  
 Best Local Similarity 100.0%; Pred. No. 9.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 437 DDDDDDDG 444  
 DB 183 DDDDDDDG 190  
 RESULT 24  
 EZH2\_MOUSE STANDARD; PRT; 746 AA.  
 ID Q61188; Q9R090;  
 AC 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Enhancer of zeste homolog 2 (ENX-1).  
 GN EZH2 OR ENX1H.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Placenta;  
 RX MEDLINE=97014262; PubMed=8661097;  
 RA Hobert O., Sures I., Closek T., Fuchs M., Ullrich A.;  
 RT "Isolation and developmental expression analysis of Enx-1, a novel  
 RT mouse Polycomb group gene.";  
 RT Mech. Dev. 55:171-184 (1996).  
 [2]  
 RN SEQUENCE OF 134-497 FROM N.A.  
 RP STRAIN=129/Sv;  
 RX MEDLINE=99160476; PubMed=10051331;  
 RA Laible G., Haynes A.R., Lebersorger A., O'Carroll D., Mattei M.G.,  
 RA Denby F., Brown S.D., Jenwein T.;  
 RT "The murine polycomb-group genes *enx1* and *enx2* map close to *hox* gene  
 RT clusters on mouse chromosomes 11 and 6.";  
 RL Mamm. Genome 10:311-314 (1999).  
 CC -1- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF GENE TRANSCRIPTION  
 AND CHROMATIN STRUCTURE.  
 CC -1- SUBUNIT: BINDS ATRX VIA THE SET DOMAIN (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ENX-1A (SHOWN HERE) AND ENX-1B;  
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED IN EARLY EMBRYOS. IN LATER  
 CC EMBRYOGENESIS, EXPRESSION RESTRICTED TO CENTRAL AND PERIPHERAL  
 CC NERVOUS SYSTEM, LIVER AND THYMUS. IN ADULT, HIGHEST EXPRESSION IN  
 CC SPLEEN, TESTIS AND PLACENTA. LOWER LEVELS IN INTESTINE AND MUSCLE  
 CC AND VERY LOW LEVELS IN BRAIN AND LIVER. NO EXPRESSION IN HEART.  
 CC THYROID GLAND, LUNG AND KIDNEY.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN BOTH ADULT AND EMBRYO WITH  
 CC HIGHEST LEVELS IN EARLY EMBRYOGENESIS.  
 CC -1- SIMILARITY: CONTAINS 1 SET DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE EZ FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U52951; AAC52655.1; -;  
 DR EMBL: AF104359; AAD54020.1; -;  
 DR MGD; MGI:107940; Ezh2.  
 DR InterPro: IPR001005; Myb\_DNA\_binding.  
 DR InterPro: IPR001214; SET.  
 DR Pfam; PF00856; SET; 1.  
 DR SMART; SM00395; SANT; 2.  
 DR SMART; SM00317; SET; 1.  
 DR PROSITE; PS50280; SET; 1.  
 KW Transcription regulation; Nuclear protein; DNA-binding;  
 KW Alternative splicing.  
 FT DOMAIN 490 495 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 FT DOMAIN 523 605 CYS-RICH.  
 FT DOMAIN 618 729 SET.  
 FT DOMAIN 724 724 DGSNNHYNYQPCDHPQPCSSCPVIAONCFKRCQCS  
 FT VARSPLIC 511 553 EC -> G (IN ISOFORM ENX-1B).  
 FT CONFLICT 159 161 MISSING (IN REF. 2).  
 SQ SEQUENCE 746 AA; 85336 MW; 0435C021963BD24E CRC64;  
 Query Match 1.8%; Score 8; DB 1; Length 746;  
 Best Local Similarity 100.0%; Pred. No. 9.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 437 DDDDDDDG 444



Db 183 DDDDDDDG 190

```

RESULT 25
CN3B_MOUSE
ID CN3B_MOUSE STANDARD; PRT; 799 AA.
AC Q61409;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE CGMP-inhibited 3',5'-cyclic phosphodiesterase B (EC 3.1.4.17) (Cyclic
  GMP inhibited phosphodiesterase B) (CGI-PDE B) (CGIPDEB) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
STRAIN=Swiss; TISSUE=Fat;
MEDLINE=97079687; PubMed=8921398;
Loebert R.W., Winterpacht A., Selpel B.U.;
Molecular cloning and chromosomal assignment of the human homologue
of the rat CGMP-inhibited phosphodiesterase 1 (PDE3A) -- a gene
involved in fat metabolism located at 11p15.1.;
Genomic 37:211-218(1996).
CC - FUNCTION: MAY PLAY A ROLE IN FAT METABOLISM.
CC - CATALYTIC ACTIVITY: Nucleoside 3',5'-cyclic phosphate + H(2)O =
  nucleoside 5'-phosphate.
CC - ENZYME REGULATION: INHIBITED BY CGMP.
CC - SUBCELLULAR LOCATION: Membrane-bound (Potential).
CC - TISSUE SPECIFICITY: ABUNDANT IN ADIPOSE TISSUES.
CC - SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE
  FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X95521; CA64775.1; .
DR MGD; MGI:133363; Pde3b.
DR InterPro; IPR003507; ME_Pplase_HDC.
DR InterPro; IPR002073; PDEase.
DR Pfam; PF00233; PDEase; 1.
DR SMART; SM00471; HDC; 1.
DR PROSITE; PS00126; PDEASE_1; 1.
DR HydroLase; CGMP; Membrane.
FT NON TER 1 1
FT DOMAIN 715 724 POLY-ASP.
FT DOMAIN 773 777 POLY-GLU.
FT NON TER 799 799
SQ SEQUENCE 799 AA; 89811 MW; B76685895F1C8B3F CRC64;

Query Match 1.8%; Score 8; DB 1; Length 799;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDDG 444
Db 716 DDDDDDDG 723

```

```

DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 95.1 kDa protein in ACT5-YCK1 intergenic region.
GN YHR131C...
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
  Du Z., Favello A., Fulton L., Gattung S., Geisler C., Kirschen J.,
  Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
  Latreille P., Louis E.J., Macri C., Marton L., Menezes S., Mouser L.,
  Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
  Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
  Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
  VIII."
RL Science 265:2077-2082(1994).
CC - SIMILARITY: TO YEAST YNL144C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U10398; AAB68414.1; .
DR PIR; S48975; S48975.
DR HSSP; Q00963; IDRO.
DR SCD; S0001173; YHR131C.
DR InterPro; IPR01849; PH.
DR Pfam; PF00169; PH; 1.
DR SMART; SM00233; PH; 1.
DR Hypothetical protein.
FT DOMAIN 312 325 POLY-ARG.
FT DOMAIN 786 840 ASP/GLU-RICH (HIGHLY ACIDIC).
SQ SEQUENCE 840 AA; 95058 MW; 3A86717D3332A0DF CRC64;

Query Match 1.8%; Score 8; DB 1; Length 840;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDDG 444
Db 829 DDDDDDDG 836

RESULT 27
DD10_HUMAN
ID DD10_HUMAN STANDARD; PRT; 875 AA.
AC Q13206;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable ATP-dependent RNA helicase DDX10 (DEAD-box protein 10).
GN DDX10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96301396; PubMed=8660968;
RA Savitsky K., Ziv Y., Bar-Shira A., Gilad S., Tagle D.A., Smith S.,
  Uziel T., Stez S., Naimias J., Sartiel A., Eddy R.L., Shovs T.B.,
  Collins F.S., Shiloh Y., Rotman G.;
RT "A human gene (DDX10) encoding a putative DEAD-box RNA helicase at
  11q22-q23."

```

Genomics 33:199-206(1996).

[2]  
 RN SEQUENCE FROM N.A.  
 RA Arai Y., Kaneko Y., Kubo T., Arai K., Hosoda F., Ohki M.,  
 RT "Molecular analysis of the chromosomal breakpoints and identification  
 of the repetitive sequences near the breakpoints of NUP98 in  
 RT therapy-related leukemia with inv(11)(p15q22)."  
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: PUTATIVE ATP-DEPENDENT RNA HELICASE.  
 CC -1- TISSUE SPECIFICITY: HIGH IN TESTIS BUT WIDELY EXPRESSED.  
 CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY. DBP4/DDX10  
 CC SUBFAMILY.  
 CC -1- DATABASE: NMBE-Atlas Genet. CytoGenet. Oncol. Haematol.;  
 CC WWW="http://www.inobioogen.fr/services/cancer/Genes/DDX10.html".

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 CC EMBL: U28042; AAC50823.1; -  
 CC EMBL: AB040537; BAB18536.1; -  
 CC SWISS-2DPAGE: Q13206; HUMAN.  
 CC Genew: HGNC:2735; DDX10.  
 CC MIM: 601235; -  
 CC InterPro: IPR001410; DEAD.  
 CC InterPro: IPR000629; DEAD\_Box.  
 CC InterPro: IPR001650; Helicase\_C.  
 CC Pfam: PF00270; DEAD; 1.  
 CC Pfam: PF00271; helicase\_C; 1.  
 CC SMART: SM00487; DEXDC; 1.  
 CC SMART: SM00490; HELIC\_C; 1.  
 CC PROSITE: PS00039; DEAD\_ATP\_HELICASE; 1.  
 CC Helicase; ATP-binding; RNA-binding.  
 CC NP BIND 113 120 ATP (POTENTIAL).  
 CC SITE 222 225 DEAD BOX.  
 CC FT CONFLICT 647 647 A -> D (IN REF. 2).  
 CC FT CONFLICT 658 658 D -> E (IN REF. 2).  
 CC FT CONFLICT 661 661 N -> K (IN REF. 2).  
 CC FT SEQUENCE 875 AA; 100815 MW; 4692EDAS6AD945B5 CRC64;

Query Match 1.8%; Score 8; DB 1; Length 875;  
 Best local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

437 DDDDDDDG 444  
 |||||  
 Db 782 DDDDDDDG 789

RESULT 28  
 ID CAPA\_HUMAN STANDARD; PRT; 938 AA.  
 AC Q3111; Q9UYB8;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Chromatin assembly factor 1 subunit A (CAF-1 subunit A) (Chromatin  
 DE assembly factor 1 p150 subunit) (CAF-I 150 kDa subunit) (CAF-Ip150).  
 GN CHAF1A OR CAF1P150 OR CAF.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Cervical adenocarcinoma;  
 RX MEDLINE=95323966; PubMed=7600578;  
 RA Kaufman P.D., Kobayashi R., Kessler N., Stillman B.;  
 RT "The p150 and p60 subunits of chromatin assembly factor I: a molecular

link between newly synthesized histones and DNA replication.";  
 Cell 81:1105-1114(1995).

[2]  
 RN SEQUENCE FROM N.A. (ISOFORMS 1, 2 AND 3).  
 RA MEDLINE=21147928; PubMed=11250073;  
 RA Dong H., Lin W., Zhang C.-K., Xiong H., Fu G., Jin W.-R., Chen R.,  
 RA Chen Z., Qi Z.-T., Huang G.M.;  
 RT "Genomic sequence and expression analyses of human chromatin assembly  
 RT factor 1 p150 gene.";  
 RT Gene 264:187-196(2001).  
 RN [3]  
 RN INTERACTION WITH PCNA, AND SUBCELLULAR LOCATION.  
 RX MEDLINE=99159824; PubMed=10052459;  
 RA Shihahara K., Stillman B.;  
 RT "Replication-dependent marking of DNA by PCNA facilitates  
 RT CAF-1-coupled inheritance of chromatin.";  
 RL Cell 96:575-585(1999).  
 RN [4]  
 RN INTERACTION WITH PCNA.  
 RX MEDLINE=20115866; PubMed=10648606;  
 RA Wogge U.G., Grandi P., Quivy J.P., Jonsson Z.O., Hubacher U.,  
 RA Becker P.B., Almouzni G.;  
 RT "A CAF-1-PCNA-mediated chromatin assembly pathway triggered by sensing  
 RT DNA damage.";  
 RL Mol. Cell. Biol. 20:1206-1218(2000).  
 RN [5]  
 RN REVIEW.  
 RX MEDLINE=20351441; PubMed=10893180;  
 RA Ridgway P., Almouzni G.;  
 RT "CAF-1 and the inheritance of chromatin states: at the crossroads of  
 RT DNA replication and repair.";  
 RL J. Cell Sci. 113:2647-2658(2000).  
 CC -1- FUNCTION: COMPLEX THAT IS THOUGHT TO MEDIATE CHROMATIN ASSEMBLY IN  
 CC DNA REPLICATION AND DNA REPAIR. ASSEMBLES HISTONE OCTAMERS ONTO  
 CC REPLICATING DNA IN VITRO. CAF-1 PERFORMS THE FIRST STEP OF THE  
 CC NUCLEOSOME ASSEMBLY PROCESS, BRINGING NEWLY SYNTHESIZED HISTONES  
 CC H3 AND H4 TO REPLICATING DNA; HISTONES H2A/H2B CAN BIND TO THIS  
 CC CHROMATIN PRECURSOR SUBSEQUENT TO DNA REPLICATION TO COMPLETE THE  
 CC HISTONE OCTAMER. P150 BINDS TO HISTONES H3 AND H4. IT MAY PLAY A  
 CC ROLE IN HETEROCHROMATIN MAINTENANCE IN PROLIFERATING CELLS BY  
 CC BRINGING NEWLY SYNTHESIZED CBX PROTEINS TO HETEROCHROMATIN DNA  
 CC REPLICATION FOCI (BY SIMILARITY).  
 CC -1- SUBUNIT: CAF-1 IS COMPOSED OF THREE SUBUNITS, P48, P60 AND P150.  
 CC P150 BINDS DIRECTLY TO P60. ONLY MINOR AMOUNTS OF P48 ARE  
 CC COMPLEXED WITH P60 AND P150 IN G1 PHASE. P150 BINDS DIRECTLY TO  
 CC PCNA AND TO CBX1.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR; DNA REPLICATION FOCI.  
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; MAY BE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- DEVELOPMENTAL STAGE: ACTIVE COMPLEX IS FOUND IN G1, S AND G2  
 CC PHASES.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 CC EMBL: U20979; AAA76736.1; -  
 CC EMBL: AF190465; AAF04291.1; -  
 CC Genew: HGNC:1910; CHAF1A.  
 CC MIM: 601246; -  
 CC DNA replication; DNA repair; Cell cycle; Chromosome; Nuclear protein;  
 CC Alternative splicing.  
 KW DOMAIN 1 296 BINDS TO PCNA.  
 FT DOMAIN 1 296 ARG/GLU/LYS-RICH.  
 FT DOMAIN 305 435 POLY-GLU.  
 FT DOMAIN 584 590 POLY-ASP.  
 FT DOMAIN 601 605 POLY-ASP.  
 FT DOMAIN 887 893 POLY-GLU.  
 FT DOMAIN 642 938 BINDS TO P60.

FT VARSPLIC 692 731 CPLETTPAGEEQTPRASKRERDEQIILQPLHNGV  
 FT VARSPLIC 732 938 -> HWHPESRDGVCRTLRVSPQSRVYINRLNSCVKSTLSC  
 FT VARSPLIC 732 938 MISSING (IN ISOFORM 2).  
 FT VARSPLIC 754 926 MISSING (IN ISOFORM 2).  
 FT VARSPLIC 757 757 T->S (IN REF. 2).  
 SQ SEQUENCE 938 AA: 105222 MW: DBD01B9F1A0E28B7 CRC64;  
 Query Match 1.8%; Score 8; DB 1; Length 938;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 434 SEGDDDD 441  
 Db 598 SEGDDDD 605

RESULT 29  
 ROR2\_HUMAN  
 ROR2\_HUMAN STANDARD: PRT: 943 AA.  
 001974; 09HAY7; 09HB61;  
 16-OCT-2001 (Rel. 40, Created)  
 15-JUN-2002 (Rel. 41, Last sequence update)  
 Tyrosine-protein kinase transmembrane receptor ROR2 precursor  
 (EC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 2).  
 GN ROR2 OR NTRK2.  
 OS Homo sapiens (Human).  
 CC Eukaryote; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 CC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=93100347; PubMed=1334494;  
 RT Maelakowski P., Carroll R.D.,  
 RT "A novel family of cell surface receptors with tyrosine kinase-like  
 RT domain".  
 RL J. Biol. Chem. 267:26181-26190(1992).  
 [2]  
 RP SEQUENCE OF 34-943 FROM N.A., AND VARIANT ILE-819.  
 RA MEDLINE=20164326; PubMed=10700182;  
 RT Oldridge M., Fortuna A.M., Maringa M., Propping P., Mansour S.,  
 RA Pollitt C., Dechiera T.M., Kimble R.B., Valenzuela D.M.,  
 RA Yancopoulos G.D., Wilkie A.O.M.,  
 RT "Dominant mutations in ROR2, encoding an orphan receptor tyrosine  
 RT kinase, cause brachydactyly type B.";  
 RL Nat. Genet. 24:275-278(2000).  
 [3]  
 RP SEQUENCE OF 34-574 FROM N.A., AND VARIANT THR-245.  
 RA MEDLINE=20442029; PubMed=10986040;  
 RT Schwabe G.C., Tinschert S., Buschow C., Meinecke P., Wolff G.,  
 RA Gilleseben-Kaebbach G., Oldridge M., Wilkie A.O.M., Koemec R.,  
 RT "Distinct mutations in the receptor tyrosine kinase gene ROR2 cause  
 RT brachydactyly type B.";  
 RL Am. J. Hum. Genet. 67:822-831(2000).  
 [4]  
 RP VARIANTS RRS C-184; W-189; W-366 AND K-620.  
 RA MEDLINE=20392394; PubMed=10932186;  
 RT Afzal A.R., Rajab A., Fenske C.D., Oldridge M., Elanko N.,  
 RA Terres-Pereira E., Teyssiez B., Murday V.A., Paton M.A.,  
 RT Wilkie A.O.M., Jeffery S.,  
 RT "Recessive Robinow syndrome, allelic to dominant brachydactyly type B,  
 RT is caused by mutation of ROR2.";  
 RL Nat. Genet. 25:419-422(2000).  
 [5]  
 RP VARIANT RRS TYR-182.  
 RA MEDLINE=20392395; PubMed=10932187;  
 RT van Bokhoven H., Celli J., Kayserili H., van Beusekom E., Balci S.,  
 RA Bruseel W., Skovby F., Kerr B., Percin E.F., Akatsu N., Brunner H.G.,  
 RT "Mutation of the gene encoding the ROR2 tyrosine kinase causes  
 RT autosomal recessive Robinow syndrome.";  
 RL Nat. Genet. 25:423-426(2000).

RN [6]  
 RP ERRATUM.  
 RA van Bokhoven H., Celli J., Kayserili H., van Beusekom E., Balci S.,  
 RA Bruseel W., Skovby F., Kerr B., Percin E.F., Akatsu N., Brunner H.G.,  
 RL Nat. Genet. 26:383-383(2000).  
 CC -1- FUNCTION: TYROSINE-PROTEIN KINASE RECEPTOR WHICH MAY BE INVOLVED  
 CC IN THE EARLY FORMATION OF THE CHONDROCYTES. IT SEEMS TO BE  
 CC REQUIRED FOR CARTILAGE AND GROWTH PLATE DEVELOPMENT.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS DURING EARLY  
 CC EMBRYONIC DEVELOPMENT. THE EXPRESSION LEVELS DROP STRONGLY AROUND  
 CC DAY 16 AND THERE ARE ONLY VERY LOW LEVELS IN ADULT TISSUES.  
 CC -1- DISEASE: DEFECTS IN ROR2 ARE A CAUSE OF BRACHYDACTYLY TYPE B  
 CC (BDB). BDB IS AN AUTOSOMAL DOMINANT SKELETAL DISORDER  
 CC CHARACTERIZED BY HYPOPLASIA/APLASIA OF DISTAL PHALANGES AND NAILS.  
 CC IN BDB THE MIDDLE PHALANGES ARE SHORT BUT IN ADDITION THE TERMINAL  
 CC PHALANGES ARE RUDIMENTARY OR ABSENT. BOTH FINGERS AND TOES ARE  
 CC AFFECTED. THE THUMBS AND BIG TOES ARE USUALLY DEFORMED.  
 CC -1- DISEASE: DEFECTS IN ROR2 ARE A CAUSE OF RECESSIVE ROBINOW SYNDROME  
 CC (RRS). RRS IS AN AUTOSOMAL DISORDER CHARACTERIZED BY SKELETAL  
 CC DYSPLASIA WITH GENERALIZED LIMB BONE SHORTENING, SEGMENTAL DEFECTS  
 CC OF THE SPINE, BRACHYDACTYLY AND A DISMORPHIC FACIAL APPEARANCE.  
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. FOR  
 CC SUBFAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 FRIZZLED (FZ) DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL; M97639; AAA60276.1; -  
 DR EMBL; AF294796; AAG01184.2; -  
 DR EMBL; AF254747; AAG01184.2; JOINED.  
 DR EMBL; AF254748; AAG01184.2; JOINED.  
 DR EMBL; AF254749; AAG01184.2; JOINED.  
 DR EMBL; AF254750; AAG01184.2; JOINED.  
 DR EMBL; AF254751; AAG01184.2; JOINED.  
 DR EMBL; AF254752; AAG01184.2; JOINED.  
 DR EMBL; AF254753; AAG01184.2; JOINED.  
 DR EMBL; AF279752; AAG33132.1; -  
 DR EMBL; AF279753; AAG33132.1; JOINED.  
 DR EMBL; AF279754; AAG33132.1; JOINED.  
 DR EMBL; AF279755; AAG33132.1; JOINED.  
 DR EMBL; AF279756; AAG33132.1; JOINED.  
 DR EMBL; AF279757; AAG33132.1; JOINED.  
 DR EMBL; AF279758; AAG33132.1; JOINED.  
 DR EMBL; AF279759; AAG33132.1; JOINED.  
 DR EMBL; AF279760; AAG33132.1; JOINED.  
 DR EMBL; AF279761; AAG33132.1; JOINED.  
 DR HSSP; P00747; IREN.  
 DR Genew; HGNC:10257; ROR2.  
 DR MIM; 602337; -  
 DR MIM; 113000; -  
 DR MIM; 268310; -  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR000024; Fz\_domain.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003598; Ig\_C2.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00047; Ig\_1.  
 DR Pfam; PF00051; kringle; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF01392; Fz\_1.  
 DR PRINTS; PR00018; KRINGLE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR ProDom; PD000395; Kringle; 1.

DR SMART; SMO0408; IGC2; 1.  
 DR SMART; SMO0130; KR; 1.  
 DR SMART; SMO0219; TYKc; 1.  
 DR PROSITE; PS50038; FZ; 1.  
 DR PROSITE; PS50021; KRINGLE\_1; FALSE\_NEG.  
 DR PROSITE; PS50070; KRINGLE\_2; 1.  
 DR PROSITE; PS50107; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS50011; PROTEIN KINASE TYR; 1.  
 DR PROSITE; PS50109; PROTEIN KINASE TYR; 1.  
 DR Tyrosinase; Tyrosine-protein kinase; ATP-binding; Receptor;  
 Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;  
 Immunoglobulin domain; Developmental protein; Polymorphism;  
 Disease mutation.  
 FT SIGNAL 1 33  
 FT CHAIN 34 943  
 FT DOMAIN 34 403  
 FT TRANSMEM 404 424  
 FT DOMAIN 425 943  
 FT DOMAIN 76 142  
 FT DOMAIN 169 303  
 FT DOMAIN 316 394  
 FT DOMAIN 473 746  
 FT DOMAIN 753 782  
 FT DOMAIN 784 857  
 FT DOMAIN 859 882  
 FT NP BIND 479 507  
 FT BINDING 507 507  
 FT ACT SITE 615 615  
 FT MOD RES 646 646  
 FT DISULFID 83 135  
 FT CARBOHYD 70 70  
 FT CARBOHYD 188 188  
 FT CARBOHYD 318 318  
 FT VARIANT 182 182  
 FT VARIANT 184 184  
 FT VARIANT 189 189  
 FT VARIANT 245 245  
 FT VARIANT 366 366  
 FT VARIANT 620 620  
 FT VARIANT 819 819  
 FT SEQUENCE 943 AA; 104726 MW; DBAC1E622B5ECA0 CRC64;  
 Query Match 1.8%; Score 8; DB 1; Length 943;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVVIME 198  
 DB 677 SYGVVIME 684

RESULT 30  
 ROR2\_MOUSE STANDARD; PRT; 944 AA.  
 ID ROR2\_MOUSE  
 AC 092138;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tyrosine-protein kinase transmembrane receptor ROR2 precursor  
 DE (EC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 2)  
 DE (ROR2).  
 DE (ROR2).  
 GN ROR2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99248426; PubMed=10231392;  
 RA Oishi T., Takeuchi S., Hashimoto R., Nagabukuro A., Ueda T., Liu Z.-J.,  
 RA Hattori T., Akira S., Matsuda Y., Yamamura H., Otani H., Minami Y.,  
 RT "Spatio-temporally regulated expression of receptor tyrosine kinases,  
 RT mror1, mror2, during mouse development: implications in development  
 RT and function of the nervous system.";  
 RL Genes Cells 4:41-56(1999).  
 [2]  
 RP CHARACTERIZATION.  
 RX MEDLINE=20164325; PubMed=10700181;  
 RA DeChiara T.M., Kimble R.B., Poueymiron W.T., Rojas J., Masiakowski P.,  
 RA Valenzuela D.M., Yancopoulos G.D.,  
 RT "Ror2, encoding a receptor-like tyrosine kinase, is required for  
 RT cartilage and growth plate development.";  
 RL Nat. Genet. 24:271-274(2000).  
 CC - FUNCTION: TYROSINE-PROTEIN KINASE RECEPTOR WHICH MAY BE INVOLVED  
 CC IN THE EARLY FORMATION OF THE CHONDROCYTES. IT SEEMS TO BE  
 CC REQUIRED FOR CARTILAGE AND GROWTH PLATE DEVELOPMENT.  
 CC - CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC - SUBCELLULAR LOCATION: Type I membrane protein.  
 CC - SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ROR  
 CC - SIMILARITY: CONTRAINS 1 KRINGLE DOMAIN.  
 CC - SIMILARITY: CONTRAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.  
 CC - SIMILARITY: CONTRAINS 1 KRINGLE DOMAIN.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL; AB010384; BAAT5481.1; --  
 DR HSSP; P00747; SHPG.  
 DR MGD; MGI:1347521; Ror2.  
 DR InterPro; IPR0000719; Euk.pkinase.  
 DR InterPro; IPR0000024; Fz.domain.  
 DR InterPro; IPR003006; Iq\_MHC.  
 DR InterPro; IPR003598; Iq\_C2.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00047; Iq; 1.  
 DR Pfam; PF00051; Kringle; 1.  
 DR Pfam; PF00063; pkinase; 1.  
 DR Pfam; PF01392; Fz; 1.  
 DR PRINTS; PR00018; KRINGLE.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk.pkinase; 1.  
 DR ProDom; PD000395; Kringle; 1.  
 DR SMART; SMO0408; IGC2; 1.  
 DR SMART; SMO0130; KR; 1.  
 DR SMART; SMO0219; TYKc; 1.  
 DR PROSITE; PS50038; FZ; 1.  
 DR PROSITE; PS50021; KRINGLE\_1; FALSE\_NEG.  
 DR PROSITE; PS50070; KRINGLE\_2; 1.  
 DR PROSITE; PS50107; PROTEIN KINASE ATP; FALSE\_NEG.  
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS50109; PROTEIN KINASE TYR; 1.  
 DR Tyrosinase; Tyrosine-protein kinase; ATP-binding; Receptor;  
 Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;  
 Immunoglobulin domain; Developmental protein.  
 FT SIGNAL 1 33  
 FT CHAIN 34 944  
 FT DOMAIN 34 403  
 FT TRANSMEM 404 424  
 FT DOMAIN 425 944  
 FT DOMAIN 425 944  
 FT DOMAIN 76 142  
 FT DOMAIN 169 303  
 FT DOMAIN 316 394  
 FT DOMAIN 473 746  
 FT DOMAIN 753 782  
 FT DOMAIN 784 857  
 FT DOMAIN 859 882  
 FT NP BIND 479 507  
 FT BINDING 507 507  
 FT ACT SITE 615 615  
 FT MOD RES 646 646  
 FT DISULFID 83 135  
 FT CARBOHYD 70 70  
 FT CARBOHYD 188 188  
 FT CARBOHYD 318 318  
 FT VARIANT 182 182  
 FT VARIANT 184 184  
 FT VARIANT 189 189  
 FT VARIANT 245 245  
 FT VARIANT 366 366  
 FT VARIANT 620 620  
 FT VARIANT 819 819  
 FT SEQUENCE 943 AA; 104726 MW; DBAC1E622B5ECA0 CRC64;

```

FT DOMAIN 76 142 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 169 303 F2.
FT DOMAIN 316 394 KRINGLE.
FT DOMAIN 473 746 PROTEIN KINASE.
FT DOMAIN 753 782 SER/THR-RICH.
FT DOMAIN 784 857 PRO-RICH.
FT NP BIND 859 882 SER/THR-RICH.
FT BINDING 479 487 ATP (BY SIMILARITY).
FT ACT SITE 507 507 ATP (BY SIMILARITY).
FT MOD RES 646 615 BY SIMILARITY.
FT DISULFID 83 135 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 70 70 BY SIMILARITY.
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 944 AA; 105050 MW; CD2EEBC710387A56 CRC64;

```

```

Query Match 1.8%; Score 8; DB 1; Length 944;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
191 SYGVVLM 198
Db 677 SYGVVLM 684

```

```

RESULT 31
VP41_YEAST
ID P38959; Q12011; P87334; STANDARD; PRT; 992 AA.
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Vacuolar assembly protein VP84 (Vacuolar morphogenesis protein VAM2).
GN VP84 OR VAM2 OR FET2 OR YDR080W OR D446 OR YD8554.13.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X2180-1A; PubMed=9111041;
RX Nakamura N., Hirata A., Ohsumi Y., Wada Y.;
RA "Vam2/VP84p and Vam6/VP83p are components of a protein complex on
RT the vacuolar membranes and involved in the vacuolar assembly in the
RT yeast Saccharomyces cerevisiae."
RT J Biol. Chem. 272:11344-11349(1997).

```

```

[2]
SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=96093910; PubMed=7483840;
RA Coster F., Joniaux J.-L., Goffeau A.;
RT "Analysis of a 32.8 kb segment of yeast chromosome IV reveals 21 open
RT reading frames, including TPS2, PPH3, RAD55, SEDI, PDC2, APR1, SSS1,
RT Slu7 and a tRNA for arginine."
RT Yeast 11:673-679(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX Richards C., Harris D.E., Barrell B.G., Rajadream M.A.;
RT Submitted (NOV-1994) to the EMBL/GenBank/DBD databases.
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=97303186; PubMed=9159129;
RA Redieky D.C., Snyder W.B., Emr S.D., Kaplan J.;
RT "Characterization of VP84, a gene required for vacuolar trafficking
RT and high-affinity iron transport in yeast."
RT Proc. Natl. Acad. Sci. U.S.A. 94:5662-5666(1997).
CC -1- FUNCTION: REQUIRED FOR VACUOLAR ASSEMBLY AND VACUOLAR TRAFFIC.
CC -1- SUBUNIT: INTERACTS WITH VP839/VAM6.
CC -1- SIMILARITY: BELONGS TO THE VP84 FAMILY.

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC EMBL; AB000223; BA119071.1;
CC EMBL; X82086; CA57607.1;
CC EMBL; Z74376; CA98899.1;
CC EMBL; Z46796; CA86802.1;
CC PIR; S48767; S48767.
CC SGD; S0002487; VP84.
CC InterPro; IPR005547; Clathrin_repeat.
CC InterPro; IPR001680; WD40.
CC Pfam; PF00400; WD40; 2.
CC Pfam; PF00637; Clathrin; 1.
CC SMART; SM00299; CLH; 1.
CC SMART; SM00320; WD40; 1.
FT DOMAIN 79 94 POLY-ASP.
FT CONFLICT 239 244 POLY-LYS.
FT CONFLICT 424 424 K -> M (IN REF. 2).
SQ SEQUENCE 992 AA; 113411 MW; E1E48B4D5A1A4005 CRC64;

```

```

Query Match 1.8%; Score 8; DB 1; Length 992;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 436 GDDDDDD 443
Db 78 GDDDDDD 85

```

```

RESULT 32
IFH1_YEAST
ID IFH1_YEAST STANDARD; PRT; 1085 AA.
AC P39520;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE IFH1 protein (IRP3 protein).
GN IFH1 OR IRP3 OR YLR223C OR L8083.9.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FL100;
RX MEDLINE=95304839; PubMed=7785326;
RA Cherel I., Thuriaux P.;
RT "The IFH1 gene product interacts with a fork head protein in
RT Saccharomyces cerevisiae."
RT Yeast 11:261-270(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Favella A., Fulcon L., Gattung S., Greco T., Kireten J.,
RA Kueba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,
RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
RA Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
RA Ricken L., Riles L., Taich A., Trevisan E., Vignati D.,
RA Wilcox L., Wohlman P., Vaudin M., Wilson R., Waterson R.;
RT Submitted (DEC-1994) to the EMBL/GenBank/DBD databases.
CC -1- FUNCTION: CONTROLS THE PRE-RNA PROCESSING MACHINERY IN
CC CONJUNCTION WITH PHL1. COULD CONVERT PHL1 FROM A REPRESSOR
CC TO AN ACTIVATOR.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; Z29488; CAA82624.1; --  
 DR EMBL; U19027; AAB67412.1; --  
 DR PIR; S47477; S47477.  
 DR SGD; S0004213; IFL1.  
 DR KMW nuclear protein; Transcription regulation.  
 FT DOMAIN 122 163 ASP/GLU-RICH (HIGHLY ACIDIC).  
 SQ SEQUENCE 1085 AA; 122491 MW; BEICDFE06213FE0 CRC64;

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 1; Length 1085;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDG 444  
 142 DDDDDDDG 149

RESULT 33  
 YAB9 YEAST STANDARD; PRT; 1131 AA.  
 AC P31380;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 128.5 kDa helicase in ATG1-TPD3 intergenic region.  
 GN YAL019W OR YAL001 OR FUN30.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX MEDLINE=93209532; PubMed=8458570;  
 RA Ouellette B.F.F., Clark M.W., Keng T., Storms R.K., Zhong M.W.,  
 Zeng B., Fortin N., Delaney S., Barton A.B., Kaback D.B., Bussey H.,  
 "Sequencing of chromosome I from Saccharomyces cerevisiae: analysis  
 of a 32 kb region between the LTEL and SP07 genes.";  
 RT Genome 36:32-42(1993).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=S288C / AB972;  
 RC MEDLINE=94193531; PubMed=8144453;  
 RX Barton A.B., Kaback D.B.;  
 RT "Molecular cloning of chromosome I DNA from Saccharomyces cerevisiae:  
 analysis of the genes in the FUN38-MAK16-SP07 region.";  
 RL J. Bacteriol. 176:1872-1880(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX MEDLINE=92221690; PubMed=1561836;  
 RA Clark M.W., Zhong M.W., Keng T., Storms R.K., Barton A.B.,  
 Kaback D.B., Bussey H.;  
 RT "Identification of a Saccharomyces cerevisiae homolog of the SNF2  
 transcriptional regulator in the DNA sequence of an 8.6 kb region in  
 the LTEL-CYS1 interval on the left arm of chromosome I.";  
 RL Yeast 8:133-145(1992).  
 CC -1- FUNCTION: MAY FUNCTION AS A NEGATIVE REGULATORY ELEMENT IN THE  
 CC PROCESS OF UV DAMAGE REPAIR.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; L05146; AAC04938.1; --  
 DR PIR; S22266; S22266.  
 DR SGD; S0000017; FUN30.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR InterPro; IPR000330; SNF2\_N.  
 DR Pfam; PF00176; SNF2\_N; 1.  
 DR Pfam; PF00271; Helicase\_C; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00490; HELIC; 1.  
 DR Hypothetical protein; Nuclear protein; DNA-binding; Helicase;  
 KW ATP-binding. 597 604 ATP (POTENTIAL).  
 FT NP BIND 703 706 DEGH BOX.  
 FT SITE 703 706  
 SQ SEQUENCE 1131 AA; 128506 MW; 7E61B9AB3A42ED2 CRC64;

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 1; Length 1131;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDD 443  
 DB 524 GDDDDDD 531

RESULT 34  
 YNF4 YEAST STANDARD; PRT; 1165 AA.  
 AC P53950;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Hypothetical 128.1 kDa protein in OMP2-WG5 intergenic region.  
 GN YNL054W OR N2467 OR YNL2467W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / FY1676;  
 RX MEDLINE=96021608; PubMed=8533472;  
 RA Berger P., Dolignon F., Crouzet M.;  
 RT "The sequence of a 4.4 kb fragment located on the left arm of  
 chromosome XIV from Saccharomyces cerevisiae.";  
 RL Yeast 11:967-974(1995).  
 RL [2]  
 RN ERRATUM.  
 RP MEDLINE=97060022; PubMed=8904343;  
 RX Berger P., Dolignon F., Crouzet M.;  
 RA Yeast 12:297-297(1996).  
 RL

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U12141; AAA9658.1; --  
 DR EMBL; Z71330; CAA95925.1; --  
 DR SGD; S0004999; YNL054W.  
 DR Hypothetical protein. 1165 AA; 128140 MW; 0414978AA3066CE7 CRC64;  
 SQ SEQUENCE 1165 AA; 128140 MW; 0414978AA3066CE7 CRC64;

Query Match  
 Best Local Similarity 1.8%; Score 8; DB 1; Length 1165;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437-DDDDDDG 444  
 DB 1085 DDDDDDDG 1092

RESULT 35  
 RNPF\_AZOV1 STANDARD; PRT; 86 AA.  
 ID RNPF\_AZOV1  
 AC QPST0;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 16-OCT-2001 (Rel. 40, Last annotation update)  
 GN Protein rnfH.  
 OS Azotobacter vinelandii.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 CC Azotobacter.  
 CC NCBI\_TaxID=354;  
 RN [1]  
 SEQUENCE FROM N.A.  
 STRAIN=DJ;  
 Rbio L.M., Rangaraj P., Roberts G.P., Ludden P.W.;  
 "Cloning and mutational analysis of the Azotobacter vinelandii gene  
 encoding the dinitrogenase gamma subunit."  
 Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 -1- SIMILARITY: BELONGS TO THE UPF0125 (RNPF) FAMILY.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL: AF302049; AAG29821.1;  
 DR InterPro: IPR005346; UPF0125.  
 DR Pfam: PF03658; UPF0125; 1.  
 FT DOMAIN 78 86 POLY-ASP.  
 SQ SEQUENCE 86 AA; 9632 MW; EBD84A7315BC25A9 CRC64;  
 -----  
 OY 437 DDDDDDD 443  
 DB 78 DDDDDDD 84  
 Query Match 1.5%; Score 7; DB 1; Length 86;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 -----  
 RESULT 36  
 KOA2\_ECOLI STANDARD; PRT; 101 AA.  
 ID KOA2\_ECOLI  
 AC P03052;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DE 01-NOV-1997 (Rel. 35, Last annotation update)  
 GN TRPB transcriptional repressor protein (Regulatory protein KORA).  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Escherichia.  
 CC NCBI\_TaxID=562;  
 RN [1]  
 SEQUENCE FROM N.A.  
 MEDLINE=84069779; PubMed=6316262;  
 RA Bechofer D.H., Figuretti D.H.;  
 RT "Map location and nucleotide sequence of kora, a key regulatory gene  
 RL of *Promesoccus plasmid RK2*."  
 Nucleic Acids Res. 11:7453-7469(1983).  
 [2]  
 SEQUENCE FROM N.A.

RX MEDLINE=86232584; PubMed=3520485;  
 RA Thomas C.M., Smith C.A.;  
 RT "The trfB region of broad host range plasmid RK2: the nucleotide  
 RT sequence reveals incC and key regulatory gene trfB/korA/korD as  
 RT overlapping genes."  
 RL Nucleic Acids Res. 14:4453-4469(1986).  
 CC -1- FUNCTION: IN CONJUNCTION WITH KORB, INHIBITS THE TRANSCRIPTION OF  
 CC KILA, TRPA AND KORB OPERONS. IN CONJUNCTION WITH KORB IS  
 CC RESPONSIBLE FOR THE NEGATIVE CONTROL OF KILC AND KILE OPERONS.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL: X03962; CA27596.1;  
 DR PIR: A03588; R6E6CK.  
 KW Plasmid; Transcription regulation; Repressor; DNA-binding.  
 FT DNA\_BIND 37 56 H-T-H MOTIF (POTENTIAL).  
 SQ SEQUENCE 101 AA; 11306 MW; 25622F8B98F21A5 CRC64;  
 -----  
 OY 243 WEADAKK 249  
 DB 89 WEADAKK 95  
 Query Match 1.5%; Score 7; DB 1; Length 101;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 -----  
 RESULT 37  
 UCRH\_YEAST STANDARD; PRT; 147 AA.  
 ID UCRH\_YEAST  
 AC P00127;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Ubiquinol-cytochrome c reductase complex 17 kDa protein (EC 1.10.2.2)  
 DE (Mitochondrial hinge protein) (Complex III polypeptide VI).  
 GN QCR6 OR UCR6 OR YFR033C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 CC NCBI\_TaxID=4932;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 28583 / FL100;  
 RC MEDLINE=84236098; PubMed=6329732;  
 RA van Loon A.P.G.M., de Groot R.J., de Haan M., Dekker A., Grivell L.A.;  
 RT "The DNA sequence of the nuclear gene coding for the 17-kd subunit VI  
 RT of the yeast ubiquinol-cytochrome c reductase: a protein with an  
 RL extremely high content of acidic amino acids."  
 EMBO J. 3:1039-1043(1984).  
 [2]  
 SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RC MEDLINE=95400292; PubMed=7670463;  
 RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,  
 RA Sasauma S.-I., Sasauma M., Tsuchiya Y., Soeda E., Yokoyama K.,  
 RA Yamazaki M., Tashiro H., Eki T.;  
 RT "Analysis of the nucleotide sequence of chromosome VI from  
 RT Saccharomyces cerevisiae."  
 RL Nat. Genet. 10:261-268(1995).  
 RN [3]  
 SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RC MEDLINE=96287654; PubMed=8686381;  
 RA Eki T., Naitou M., Hagiwara H., Abe M., Ozawa M., Sasauma S.-I.,  
 RA Sasauma M., Tsuchiya Y., Shibata T., Watanabe K., Ono A.,  
 RA Yamazaki M.-A., Tashiro H., Hanaka F., Murakami Y.;

RT "Fifteen open reading frames in a 30.8 kb region of the right arm of  
 RT chromosome VI from *Saccharomyces cerevisiae*." ;  
 RL Yeast 12:1177-190(1996).  
 CC -1- FUNCTION: THIS IS A COMPONENT OF THE UBIQUITINOL-CYTOCHROME C  
 CC REDUCTASE COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH  
 CC IS PART OF THE MITOCHONDRIAL RESPIRATORY CHAIN. THIS PROTEIN MAY  
 CC MEDATE FORMATION OF THE COMPLEX BETWEEN CYTOCHROMES C AND C1.  
 CC -1- FUNCTION: THIS HIGHLY ACIDIC PROTEIN MAY MEDATE THE FORMATION OF  
 CC THE COMPLEX BETWEEN CYTOCHROMES C AND C1.  
 CC -1- CATALYTIC ACTIVITY: OH(2) + 2 ferri-cytochrome c = Q + 2  
 CC ferri-cytochrome c.  
 CC SUBUNIT: FUNG1 BCL COMPLEX CONTAINS 10 SUBUNITS; 3 RESPIRATORY  
 CC SUBUNITS, 2 CORE PROTEINS AND 5 LOW-MOLECULAR WEIGHT PROTEINS.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial inner membrane.  
 CC -1- SIMILARITY: BELONGS TO THE UOCRH/QCR6 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X00551; CAA5220.1; -;  
 DR EMBL; D50617; BAA09272.1; -;  
 DR PIR; A00120; RDBYUC.  
 DR SGD; S0001929; QCR6.  
 DR Mitochondrion; Inner membrane; Electron transport; Respiratory chain;  
 KW Oxidoreductase.  
 FT DOMAIN 26 80 ASP/GLU-RICH (ACIDIC).  
 FT CONFLICT 2 2 G -> D (IN REF. 1).  
 SQ SEQUENCE 147 AA; 17257 MW; 148B9C0EB57BE8F CRC64;  
 QY Query Match 1.5%; Score 7; DB 1; Length 147;  
 QY Best Local Similarity 100.0%; Pred. No. 22;  
 QY Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 437 DDDDDD 443  
 Db 59 DDDDDD 65

RL Science 269:496-512(1995).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).  
 RX MEDLINE=99182411; PubMed=10080886;  
 RA Henig M., Dale G.E., D'Arcy A., Daniel F., Fischer S., Gray C.P.,  
 RA Toland S., Mueller F., Page M.G.P., Pattison P., Oshner C.;  
 RT "The structure and function of the 6-hydroxymethyl-7,8-dihydropterin  
 RT pyrophosphokinase from *Haemophilus influenzae*." ;  
 RL J. Mol. Biol. 287:211-219(1999).  
 CC -1- CATALYTIC ACTIVITY: ATP + 2-amino-4-hydroxy-6-hydroxymethyl-7,8-  
 CC dihydropteridine = AMP + 2-amino-7,8-dihydro-4-hydroxy-6-  
 CC (4-phosphooxymethyl)pteridine.  
 CC -1- PATHWAY: dihydrofolate biosynthesis; first step.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- SIMILARITY: BELONGS TO THE HPPK FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U32691; AAC21742.1; -;  
 DR PDB; 1CBK; 01-MAR-00.  
 DR TIGR; H10064; -;  
 DR InterPro; IPR000550; Hppk.  
 DR Pfam; PF01288; HPPK; 1.  
 DR PROSITE; PS00794; HPPK; 1.  
 KW Transferrase; Folate biosynthesis; 3D-structure; Complete proteome.  
 SQ SEQUENCE 160 AA; 18299 MW; 136CD15F8844FDD CRC64;  
 QY Query Match 1.5%; Score 7; DB 1; Length 160;  
 QY Best Local Similarity 100.0%; Pred. No. 24;  
 QY Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 222 NERLTP 228  
 Db 108 NERLTP 114

RESULT 38  
 HPPK\_HAEIN  
 ID HPPK\_HAEIN STANDARD; PRT; 160 AA.  
 AC P43777;  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 2-amino-4-hydroxy-6-hydroxymethyl-dihydropteridine pyrophosphokinase  
 DE (EC 2.7.6.3) (7,8-dihydro-6-hydroxymethylpterin-pyrophosphokinase)  
 DE (HPPK) (6-hydroxymethyl-7,8-dihydropterin pyrophosphokinase) (PPPK).  
 GN FOLK OR H10064.  
 OS *Haemophilus influenzae*.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC *Haemophilus*.  
 OC NCBI\_TaxID=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=Rd / KM20 / ATCC 51907;  
 RX MEDLINE=9535630; PubMed=7542800;  
 RA Fleisemann A.R., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom B., Cotton M.D.,  
 RA Uetrich T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of *Haemophilus influenzae*  
 RT Rd." ;

RESULT 39  
 UBC2\_YEAST  
 ID UBC2\_YEAST STANDARD; PRT; 172 AA.  
 AC P06104;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Ubiquitin-conjugating enzyme E2-20 kDa (EC 6.3.2.19)  
 DE (Ubiquitin-protein ligase) (Ubiquitin carrier protein).  
 GN UBC2 OR RAD6 OR YGL058W.  
 OS *Saccharomyces cerevisiae* (Baker's yeast).  
 OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OC NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP MEDLINE=85113143; PubMed=3881753;  
 RX Reynolds P., Weber S., Prakash L.;  
 RT "RAD6 gene of *Saccharomyces cerevisiae* encodes a protein containing a  
 RT tract of 13 consecutive aspartates." ;  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:168-172(1985).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=S288C;  
 RX MEDLINE=9737793; PubMed=9234674;  
 RA Feuerhahn M., de Montigny J., Potier S., Souciet J.-L.;  
 RT "The characterization of two new clusters of duplicated genes  
 RT suggests a 'Lego' organization of the yeast *Saccharomyces cerevisiae*  
 RT chromosomes." ;  
 RL Yeast 13:861-869(1997).  
 RN [3]



RP SEQUENCE OF 77-91, AND FUNCTION.  
 RX MEDLINE=87315384; PubMed=3306404;  
 RA Jenech S., McGrath J.P., Varshavsky A.;  
 RT "The yeast DNA repair gene RAD6 encodes a ubiquitin-conjugating  
 RL Nature 329:131-134(1987).  
 [4]  
 RP MUTAGENESIS OF CYS-88.  
 RX MEDLINE=90207263; PubMed=2157209;  
 RA Sung P., Prakash S., Prakash L.;  
 RT "Mutation of cysteine-88 in the Saccharomyces cerevisiae RAD6 protein  
 RT abolishes its ubiquitin-conjugating activity and its various  
 RL biological functions.";  
 RN Proc. Natl. Acad. Sci. U.S.A. 87:2695-2699(1990).  
 [5]  
 RP REQUIREMENT FOR E3.  
 RX MEDLINE=91293094; PubMed=2065660;  
 RA Sung P., Berleth E., Pickart C., Prakash S., Prakash L.;  
 RT "Yeast Rad6 encoded ubiquitin conjugating enzyme mediates protein  
 RT degradation dependent on the N-end-recognizing E3 enzyme.";  
 RN EMBO J. 10:2187-2193(1991).  
 [6]  
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).  
 RX MEDLINE=98165808; PubMed=947353;  
 RA Worthylake D.K., Prakash S., Prakash L., Hill C.P.;  
 RT "Crystal structure of the Saccharomyces cerevisiae ubiquitin-  
 RT conjugating enzyme Rad6 at 2.6-A resolution.";  
 RL J. Biol. Chem. 273:6271-6276(1998).  
 [7]  
 RP FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO  
 CC OTHER PROTEINS. UBC2 IS ACTIVE ON HISTONES. IT IS REQUIRED FOR  
 CC POSTREPLICATION REPAIR OF UV-DAMAGED DNA AND SPOULATION. UBC2  
 CC CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +  
 CC diphosphate + protein N-ubiquityllysine.  
 CC SUBUNIT: Interacts with Rad18.  
 CC SUBCELLULAR LOCATION: Nuclear.  
 CC DOMAIN: THE ACIDIC-TAIL DOMAIN OF RAD6 IS IMPORTANT FOR  
 CC POLYUBQUITINATION OF HISTONES. AS WELL AS FOR SPOULATION.  
 CC MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR  
 CC UBIQUITIN-THIOLESTER FORMATION.  
 CC SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: K02962; AAA34952.1; -  
 CC EMBL: 272580; CAA6761.1; -  
 CC PIR: A21906; A21906.  
 CC PDB: 1AYZ; 26-AUG-98.  
 CC SGD: S0003026; RAD6.  
 CC InterPro: IPR00608; UBC\_conjugat.  
 CC Pfam: PF00179; UQ\_con; 1.  
 CC ProDom: PD000461; UBC\_conjugat; 1.  
 CC SMART: SM00212; UBC; 1.  
 CC PROSITE: PS00183; UBIQUITIN\_CONJUGAT\_1; 1.  
 CC PROSITE: PS0127; UBIQUITIN\_CONJUGAT\_2; 1.  
 CC Ubl conjugation pathway; ligase; DNA repair; Sporulation;  
 CC Nuclear protein; Multigene family; 3D-structure.  
 CC FT DOMAIN 150 172 ASP/GLU-RICH (ACIDIC TAIL).  
 CC MOD\_RES 1 1 BLOCKED.  
 CC BINDING 88 88 UBIQUITIN.  
 CC FT MUTAGEN 88 88 C-2A V. LOSS OF ACTIVITY.  
 CC SEQUENCE 172 AA; 19705 MW; 5F56DCC28ABBD60F CRC64;  
 CC  
 CC Query Match 1.5%; Score 7; DB 1; Length 172;  
 CC Best Local Similarity 100.0%; Pred. No. 26;  
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
 DB 157 DDDDDDD 163  
 RESULT 40  
 ID NUM6 BRABR STANDARD; PRT; 173 AA.  
 AC P43194;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE NADH-ubiquinone oxidoreductase chain 6 (EC 1.6.5.3).  
 GN MTND6 OR ND6 OR NADH6.  
 OS Brachyramphus brevirostris (Kiltitz murrelet).  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Archoseura; Aves; Neognathae; Charadriiformes; Alcidae;  
 OC Brachyramphus.  
 OX NCBI\_TaxID=28693;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94363783; PubMed=8082208;  
 RA Moum T., Willaen N.P., Johansen S.;  
 RT "Intragenic rearrangements in the mitochondrial NADH dehydrogenase  
 RT subunit 6 gene of vertebrates.";  
 RL Curr. Genet. 25:554-557(1994).  
 [2]  
 RP CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: X73922; CAA52127.1; -  
 CC InterPro: IPR001457; Oxidored\_q3.  
 CC Pfam: PF00499; oxidored\_q3; 1.  
 CC K02962; NAD; Ubiquinone; Mitochondrion.  
 CC QY SEQUENCE 173 AA; 18334 MW; 307D4AD6AC8C419 CRC64;  
 CC  
 CC Query Match 1.5%; Score 7; DB 1; Length 173;  
 CC Best Local Similarity 100.0%; Pred. No. 26;  
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 3 SLGASV 9  
 CC DB 45 SLGASV 51  
 CC  
 CC RESULT 41  
 CC ID NUM6 BRAMA STANDARD; PRT; 173 AA.  
 CC AC P43195;  
 CC DT 01-NOV-1995 (Rel. 32, Created)  
 CC DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 CC DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 CC DE NADH-ubiquinone oxidoreductase chain 6 (EC 1.6.5.3).  
 CC GN MTND6 OR ND6 OR NADH6.  
 CC OS Brachyramphus marmoratus (Marbled murrelet).  
 CC Mitochondrion.  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Archoseura; Aves; Neognathae; Charadriiformes; Alcidae;  
 CC Brachyramphus.  
 CC OX NCBI\_TaxID=28694;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RX MEDLINE=94363783; PubMed=8082208;  
 CC RA Moum T., Willaen N.P., Johansen S.;  
 CC "Intragenic rearrangements in the mitochondrial NADH dehydrogenase

RT subunit 6 gene of vertebrates."  
 RL Curr. Genet. 25:554-557(1994).  
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X73923; CA52128.1; -  
 DR InterPro; IPR001457; Oxidored\_q3.  
 DR Pfam; PF00499; Oxidored\_q3; 1.  
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
 SQ SEQUENCE 173 AA; 18444 MW; F4065804ADP85499 CRC64;  
 Query Match 1.5%; Score 7; DB 1; Length 173;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 3 SLGASFV 9  
 DB 45 SLGASFV 51  
 RESULT 42  
 VIRR AGRTU STANDARD; PRT; 175 AA.  
 ID VIRR AGRTU  
 AC 052278;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE VIRR protein.  
 GN  
 OS Agrobacterium tumefaciens.  
 OG PlasmidPT1A6NC.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 NCBI\_TaxID=358;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A6NC;  
 RA Zhu J., Winans S.C.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE UPF0066 (VIRR) FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF242881; AAB91563.1; -  
 DR InterPro; IPR001378; UPF0066.  
 DR Pfam; PF01980; UPF0066; 1.  
 DR ProDom; PD006705; UPF0066; 1.  
 DR TIGRFAMs; TIGR00104; UPF0066; 1.  
 DR PROSITE; PS01318; UPF0066; 1.  
 KW Plasmid.  
 SQ SEQUENCE 175 AA; 18820 MW; 18AF20BB61B087F CRC64;  
 Query Match 1.5%; Score 7; DB 1; Length 175;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 392 AGAVVMS 398  
 DB 9 AGAVVMS 15

RESULT 43  
 NO29 XENLA STANDARD; PRT; 183 AA.  
 ID NO29 XENLA  
 AC 042584;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Nucleoplasmalmin-like protein NO29 (NOVA).  
 OS Xenopus laevis (African clawed frog)  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;  
 CC Xenopodinae; Xenopus.  
 NCBI\_TaxID=8355;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=97470976; PubMed=9326619;  
 RA Zivres R.R.F., Schmidt-Zachmann M.S., Franke W.W.;  
 RT "Identification of a small, very acidic constitutive nucleolar protein  
 (NO29) as a member of the nucleoplasmalmin family."  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:11387-11392(1997).  
 CC -1- SUBCELLULAR LOCATION: Nucleolus; nucleolar.  
 CC -1- SIMILARITY: BELONGS TO THE NUCLEOPLASMIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; Z65983; CAB06652.1; -  
 DR InterPro; IPR004301; Nucleoplasmalmin.  
 DR Pfam; PF03066; Nucleoplasmalmin; 1.  
 KW Nuclear protein.  
 FT DOMAIN 127 165 ASP/GLU-RICH (HIGHLY ACIDIC).  
 FT DOMAIN 135 145 POLY-ASP.  
 FT DOMAIN 146 162  
 FT DOMAIN 162  
 SQ SEQUENCE 183 AA; 20121 MW; C1F5DE503FCETB9F CRC64;  
 Query Match 1.5%; Score 7; DB 1; Length 183;  
 Best Local Similarity 100.0%; Pred. No. 27;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 437 DDDDDDD 443  
 DB 146 DDDDDDD 152  
 RESULT 44  
 HMGT ONCMY STANDARD; PRT; 204 AA.  
 ID HMGT ONCMY  
 AC P07746; G91200;  
 DT 01-MAR-1988 (Rel. 08, Created)  
 DT 01-MAR-1988 (Rel. 10, Last sequence update)  
 DT 15-JUN-1999 (Rel. 38, Last annotation update)  
 DE High mobility group-T protein (HMG-T) (HMG-1).  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 CC Proteanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 NCBI\_TaxID=8022;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85269614; PubMed=4022777;  
 RA Penteoet B.T., Wright J.M., Dixon G.H.;  
 RT "Isolation and sequence of cDNA clones coding for a member of the  
 family of high mobility group proteins (HMG-T) in trout and analysis  
 of HMG-T-mRNA's in trout tissues."  
 RL Nucleic Acids Res. 13:4871-4888(1985).  
 RN [2]

RA REVISIONS.  
 RA Wright J.;  
 RL Submitted (MAY-1986) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=95045507; PubMed=7957172;  
 RA Sctos M., Nishikawa S., Dixon G.H.;  
 RT "CDNA sequence and structure of a gene encoding trout testis high-  
 RL mobility-group-1 protein";  
 CC Eur. J. Biochem. 225:581-591(1994).  
 CC -1- FUNCTION: BINDS PREFERENTIALLY SINGLE-STRANDED DNA AND UNWINDS  
 CC DOUBLE STRANDED DNA.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: BELONGS TO THE HMG1/HMG2 PROTEIN FAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 HMG BOXES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X02666; CAA26500.1; -;  
 DR EMBL; L32859; AAA58771.1; -;  
 DR PIR; A24019; A24019.  
 DR HSSP; P07155; IAAB.  
 DR InterPro; IPR000910; HMG\_12 box.  
 DR InterPro; IPR000135; HgHmblty\_12.  
 DR Pfam; PF00505; HMG box; 2.  
 DR PRINTS; PR00886; HIGHMOBILITY12.  
 DR SMART; SM00398; HMG; 2.  
 DR PROSITE; PS00353; HMG1\_2; 1.  
 DR Nuclear protein; Chromosomal  
 KM DNA BIND 8 78 protein; DNA-binding; Repeat.  
 FT DNA\_BIND 9 162 HMG BOX 1.  
 FT DOMAIN 181 204 ASP/GLU-RICH (ACIDIC).  
 FT CONFLICT 22 22 R -> C (IN REF. 3).  
 FT CONFLICT 149 149 R -> K (IN REF. 3).  
 SQ SEQUENCE 204 AA; 23572 MW; 6A858A144AD69A0 CRC64;  
 Query Match 1.5%; Score 7; DB 1; Length 204;  
 Best Local Similarity 100.0%; Pred. No. 30;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 437 DDDDDD 443  
 181 DDDDDD 187  
 RESULT 45  
 HIT\_MOUSE STANDARD; PRT; 207 AA.  
 ID HIT\_MOUSE  
 AC Q07133;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Histone H1c (Testicular H1 histone).  
 GN H1PT OR H1T OR H1F3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Leukocyte;  
 RX MEDLINE=94060108; PubMed=8241275;  
 RA Drabent B., Bode C., Doenecke D.;  
 RT "Structure and expression of the mouse testicular H1 histone gene  
 RT (H1c).";  
 RL Blochim. Biophys. Acta 1216:311-313(1993).

RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CD-1; TISSUE=Testis;  
 RA van Wert J., Wright J., Wolfe S.A., Grimes S.R.;  
 RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: HISTONES H1 ARE NECESSARY FOR THE CONDENSATION OF  
 CC NUCLEOSOME CHAINS INTO HIGHER ORDER STRUCTURES.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- DEVELOPMENTAL STAGE: THIS HISTONE IS A TESTIS-SPECIFIC H1 VARIANT  
 CC THAT APPEARS DURING MEIOSIS IN SPERMATOGENESIS.  
 CC -1- SIMILARITY: BELONGS TO THE HISTONE H1/H5 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L28753; AAB38417.1; -;  
 DR EMBL; X72805; CAA51325.1; -;  
 DR EMBL; U06232; AAA18359.1; -;  
 DR PIR; S33226; S33226.  
 DR PIR; S43434; S43434.  
 DR HSSP; P08287; IGHC.  
 DR MGD; MGI:107502; H1F3.  
 DR InterPro; IPR001386; Histone\_H1/H5.  
 DR Pfam; PF00538; linker histone; 1.  
 DR ProDom; PD000373; Linkerhist\_N; 1.  
 DR SMART; SM00526; H1S; 1.  
 DR Chromosomal protein; Nuclear protein; DNA-binding; Multigene family;  
 KM Acetylation; Spermatogenesis; Testis.  
 FT INIT\_MET 0 0  
 FT MOD\_RES 1 1 ACETYLATION (BY SIMILARITY).  
 FT DOMAIN 37 110 GLOBULAR.  
 FT CONFLICT 163 163 R -> G (IN REF. 1).  
 SQ SEQUENCE 207 AA; 21508 MW; A3C847CA97976C44 CRC64;  
 Query Match 1.5%; Score 7; DB 1; Length 207;  
 Best Local Similarity 100.0%; Pred. No. 31;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 415 SKRGRK 421  
 Db 21 SKRGRK 27  
 RESULT 46  
 EFLB\_ORYSA STANDARD; PRT; 228 AA.  
 ID EFLB\_ORYSA  
 AC Q40680;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Elongation factor 1-beta (EF-1-beta).  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriatropidae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Hayayuki; TISSUE=Anther;  
 RX MEDLINE=94139885; PubMed=8307147;  
 RA Matsunoto S., Terui Y., Shikong X., Taira H., Ejiri S.I.;  
 RT "Cloning and characterization of the cDNA encoding rice elongation  
 RT factor 1 beta.";  
 RL FEBS Lett. 338:103-106(1994).  
 CC -1- FUNCTION: EF-1-BETA AND EF-1-BETA' STIMULATE THE EXCHANGE OF  
 CC GDP BOUND TO EF-1-ALPHA TO GTP.  
 CC -1- SUBUNIT: EF-1 IS COMPOSED OF FOUR SUBUNITS: ALPHA, BETA,  
 CC BETA', AND GAMMA.

CC -1- SIMILARITY: BELONGS TO THE EF-1-BETA/EF-1-DELTA FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; D23674; BAA04903.1; -  
 CC InterPro: IPR001326; EF1\_BD.  
 CC Pfam: PF00736; EF1BD; 1.  
 CC PROSITE; PS00824; EF1BD\_1; 1.  
 CC PROSITE; PS00825; EF1BD\_2; 1.  
 CC Elongation factor; Protein biosynthesis.  
 CC INT MET 0 BY SIMILARITY.  
 CC SEQUENCE 228 AA; 24731 MW; D3CE96C9A7B4E9FD CRC64;  
 SQ  
 Query Match 1.5%; Score 7; DB 1; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 437 DDDDDDD 443  
 Db 100 DDDDDD 106  
 RESULT 47  
 EF1B BETVU STANDARD; PRT; 230 AA.  
 ID EF1B BETVU  
 AC 081918;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DT Elongation factor 1-beta (EF-1-beta).  
 DE Elongation factor 1-beta (EF-1-beta).  
 OS Beta vulgaris (Sugar beet).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 CC Caryophyllales; Caryophyllales; Chenopodiaceae; Beta.  
 CC NCBI\_TaxID=3555;  
 RX  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=CV. INZCHTLNIE KMS VV-D/ZR5; TISSUE=leaf;  
 RA Viereck R.;  
 RT "Nucleotide sequence of the elongation factor 1-beta from Beta  
 vulgaris";  
 Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 -1- FUNCTION: EF-1-BETA AND EF-1-BETA STIMULATE THE EXCHANGE OF  
 GDP BOUND TO EF-1-ALPHA TO GTP.  
 -1- SUBUNIT: EF-1 IS COMPOSED OF FOUR SUBUNITS: ALPHA, BETA,  
 BETA', AND GAMMA.  
 -1- SIMILARITY: BELONGS TO THE EF-1-BETA/EF-1-DELTA FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Z97067; CAB09803.1; -  
 CC InterPro: IPR001326; EF1\_BD.  
 CC Pfam; PF00736; EF1BD; 1.  
 CC PROSITE; PS00824; EF1BD\_1; 1.  
 CC PROSITE; PS00825; EF1BD\_2; 1.  
 CC Elongation factor; Protein biosynthesis.  
 CC INT MET 0 BY SIMILARITY.  
 CC SEQUENCE 230 AA; 24545 MW; EBF120B8FEC41EAO CRC64;  
 SQ  
 Query Match 1.5%; Score 7; DB 1; Length 230;  
 Best Local Similarity 100.0%; Pred. No. 34;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 437 DDDDDDD 443  
 Db 101 DDDDDD 107  
 RESULT 48  
 ICP3\_HSVIN STANDARD; PRT; 245 AA.  
 ID ICP3\_HSVIN  
 AC P37319;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE Infected cell protein ICP34.5 (Neurovirulence factor ICP34.5).  
 GN ICP34.5.  
 OS Herpes simplex virus (type 1 / strain MOH-10).  
 CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 CC Alphaherpesvirinae; Simplexvirus.  
 CC NCBI\_TaxID=37107;  
 RX  
 RN SEQUENCE FROM N.A.  
 RC MEDLINE=90156494; PubMed=2154589;  
 RA Chou J., Roizman B.;  
 RT "The herpes simplex virus 1 gene for ICP34.5, which maps in inverted  
 RT repeats, is conserved in several limited-passage isolates but not in  
 RT strain 17syn+";  
 RL J. Virol. 64:1014-1020(1990).  
 CC -1- FUNCTION: NOT KNOWN; HAS A PROFOUND EFFECT ON THE GROWTH OF THE  
 CC VIRUS IN CENTRAL NERVOUS SYSTEM TISSUE.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; IN INFECTED CELL.  
 CC -1- DEVELOPMENTAL STAGE: ACCUMULATES LATE IN INFECTION.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M33700; AAA45791.1; -  
 CC Repeat; Late protein.  
 KW Repeat; Late protein.  
 FT DOMAIN 155 172  
 SQ SEQUENCE 245 AA; 25981 MW; 55ADA50AB1B8CA11 CRC64;  
 Query Match 1.5%; Score 7; DB 1; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 437 DDDDDDD 443  
 Db 69 DDDDDD 75  
 RESULT 49  
 ICP3\_HSV1 STANDARD; PRT; 248 AA.  
 ID ICP3\_HSV1  
 AC P36313;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Infected cell protein ICP34.5 (Neurovirulence factor ICP34.5).  
 GN ICP34.5.  
 OS Herpes simplex virus (type 1 / strain 17).  
 CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 CC Alphaherpesvirinae; Simplexvirus.  
 CC NCBI\_TaxID=10299;  
 RX  
 RN SEQUENCE FROM N.A.  
 RC MEDLINE=88274327; PubMed=2839594;  
 RA McGeech D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,  
 RA McNab D., Perry L.J., Scott J.E., Taylor P.;

```

RT RT "The complete DNA sequence of the long unique region in the genome of
RT RT herpes simplex virus type 1."
RT J. Gen. Virol. 69:1531-1574 (1988).
RT [2]
RT REVISIONS.
RT MEDLINE=92341080; PubMed=1321882;
RT Dolan A., McKie E., McLean A.R., McGeoch D.J.,
RT "Status of the ICP34.5 gene in herpes simplex virus type 1 strain
RT 17."
RT J. Gen. Virol. 73:971-973 (1992).
RT -1- FUNCTION: NOT KNOWN, HAS A PROFOUND EFFECT ON THE GROWTH OF THE
RT VIRUS IN CENTRAL NERVOUS SYSTEM TISSUE.
RT -1- SUBCELLULAR LOCATION: CYTOPLASMIC, IN INFECTED CELL.
RT -1- DEVELOPMENTAL STAGE: ACCUMULATES LATE IN INFECTION.
RT CC This SWISS-PROT entry is copyright. It is produced through a collaboration
RT CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
RT CC the European Bioinformatics Institute. There are no restrictions on its
RT CC use by non-profit institutions as long as its content is in no way
RT CC modified and this statement is not removed. Usage by and for commercial
RT CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
RT CC or send an email to license@isb-sib.ch.)
RT CC -----
DR EMBL, X14112; NOT_ANNOTATED_CDS.
DR PIR, JQ1682; JQ1682.
DR KW Repeat.
DR FT DOMAIN
DR SQ SEQUENCE 161 175 5 X 3 AA TANDEM REPEATS OF A-T-P.
DR 248 AA; 26185 MW; AD334AF0C9CABA65 CRC64;
DR
DR Query Match
DR Best Local Similarity 1.5%; Score 7; DB 1; Length 248;
DR Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DR
DR 437 DDDDDDD 443
DR 75 DDDDDDD 81
DR
DR RESULT 50
DR YB95_ARATH STANDARD; PRT; 250 AA.
DR ID YB95_ARATH
DR AC O8W487;
DR DT 15-JUN-2002 (Rel. 41, Created)
DR DT 15-JUN-2002 (Rel. 41, Last sequence update)
DR DT 15-JUN-2002 (Rel. 41, Last annotation update)
DR DE Hypothetical protein At2g39795, mitochondrial precursor.
DR GN At2g39795 OR T517.20.
DR GN Arabidopsis thaliana (Mouse-ear cress).
DR CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
DR CC Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
DR CC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
DR CC NCB1_TaxID=3702;
DR CC [1]
DR CC SEQUENCE FROM N.A.
DR RP STRAIN=cv. Columbia;
DR RX MEDLINE=20083487; PubMed=10617197;
DR Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
DR Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
DR Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
DR Mottet K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Unayam L.,
DR Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Cready T.H.,
DR Goodman H.M., Somerville C.R., Coppenhaver G.P., Preuss D.,
DR Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
DR Venter J.C.;
DR "Sequence and analysis of chromosome 2 of the plant Arabidopsis
DR thaliana."
DR Nature 402:761-768 (1999).
DR
DR RP SEQUENCE FROM N.A.
DR RP STRAIN=cv. Columbia;
DR RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Riken Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGECC).";

```

```

RT Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
RT -1- SUBCELLULAR LOCATION: Mitochondrial matrix (Potential).
RT -1- SIMILARITY: BELONGS TO THE MAM33 FAMILY.
RT CC
RT CC This SWISS-PROT entry is copyright. It is produced through a collaboration
RT CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
RT CC the European Bioinformatics Institute. There are no restrictions on its
RT CC use by non-profit institutions as long as its content is in no way
RT CC modified and this statement is not removed. Usage by and for commercial
RT CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
RT CC or send an email to license@isb-sib.ch.)
RT CC -----
DR EMBL, AC003000; AAM14855.1;
DR DR EMBL, AY062757; AALJ2835.1;
DR DR InterPro; IPR003428; MAM33.
DR DR Pfam; PF02330; MAM33; 1.
DR FT TRANSIT
DR FT CHAIN 1 250 MITOCHONDRION (POTENTIAL).
DR FT SEQUENCE 250 AA; 28061 MW; EB8884B390FE7D47 CRC64;
DR
DR Query Match
DR Best Local Similarity 1.5%; Score 7; DB 1; Length 250;
DR Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DR
DR 437 DDDDDDD 443
DR 132 DDDDDDD 138
DR
DR RESULT 51
DR ICP3_HSV1D STANDARD; PRT; 252 AA.
DR ID ICP3_HSV1D
DR AC P37318;
DR DT 01-OCT-1994 (Rel. 30, Created)
DR DT 01-OCT-1994 (Rel. 30, Last sequence update)
DR DT 01-OCT-1994 (Rel. 30, Last annotation update)
DR DE Infected cell protein ICP34.5 (Neurovirulence factor ICP34.5).
DR GN ICP34.5.
DR OS Herpes simplex virus (type 1 / strain CVG-2).
DR CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
DR CC Alphaherpesvirinae; Simplexvirus.
DR CC NCB1_TaxID=37106;
DR CC [1]
DR CC SEQUENCE FROM N.A.
DR RP MEDLINE=90156494; PubMed=2154589;
DR RA Chou J., Roizman B.;
RT "The herpes simplex virus 1 gene for ICP34.5, which maps in inverted
RT repeats, is conserved in several limited-passage isolates but not in
RT strain 17syn+."
RT J. Virol. 64:1014-1020 (1990).
RT CC -1- FUNCTION: NOT KNOWN, HAS A PROFOUND EFFECT ON THE GROWTH OF THE
RT CC VIRUS IN CENTRAL NERVOUS SYSTEM TISSUE.
RT CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, IN INFECTED CELL.
RT CC -1- DEVELOPMENTAL STAGE: ACCUMULATES LATE IN INFECTION.
RT CC
RT CC This SWISS-PROT entry is copyright. It is produced through a collaboration
RT CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
RT CC the European Bioinformatics Institute. There are no restrictions on its
RT CC use by non-profit institutions as long as its content is in no way
RT CC modified and this statement is not removed. Usage by and for commercial
RT CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
RT CC or send an email to license@isb-sib.ch.)
RT CC -----
DR EMBL, M33701; AAA45792.1;
DR DR Repeat; late protein.
DR KW Repeat; late protein.
DR FT DOMAIN
DR SQ SEQUENCE 252 AA; 26601 MW; A6ED049FEAA79A CRC64;
DR
DR Query Match
DR Best Local Similarity 1.5%; Score 7; DB 1; Length 252;
DR Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 437 DDDDDDD 443  
 Db 76 DDDDDDD 82

RESULT 52  
 ASP\_PLAFS STANDARD; PRT; 253 AA.

AC P13825;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Aspartic acid-rich protein precursor.  
 OS Plasmodium falciparum (isolate Iem7 / Senegal).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.  
 NX NCBI\_TaxID=5845;

SEQUENCE FROM N.A. PubMed=303811;  
 MEDLINE=87270765; Andrieu B., le Bras J., Gallibert F.,  
 Lenstra R., D'Auriol L.,  
 "Cloning and sequencing of Plasmodium falciparum DNA fragments  
 containing repetitive regions potentially coding for histidine-rich  
 proteins: identification of two overlapping reading frames."  
 Biochem. Biophys. Res. Commun. 146:368-377(1987).  
 RT Biochem. Biophys. Res. Commun. 146:368-377(1987).  
 RL Biochem. Biophys. Res. Commun. 146:368-377(1987).  
 CC -1- MISCELLANEOUS: THIS PROTEIN IS CODED ON THE REVERSE STRAND OF AN  
 HISTIDINE-RICH PROTEIN.  
 CC -1- SIMILARITY: SOME SIMILARITY WITH HUMAN SET PROTEIN.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M17028; AAA9620.1; -  
 DR PIR; B29653; B29653.  
 DR InterPro; IPR002164; NAF\_family.  
 DR Pfam; PF00956; NAF\_1.  
 DR Malaria; Antigen; Signal.  
 KW Malaria; Antigen; Signal.  
 FT SIGNAL 1 22  
 FT CHAIN 23 253  
 FT DOMAIN 204 253  
 FT SEQUENCE 253 AA; 30248 MW; 83E585DE074B8504 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDD 443  
 Db 218 DDDDDDD 224

RESULT 53  
 ID L139 CAEEL STANDARD; PRT; 253 AA.

AC P34684;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Homeobox protein 11n-39.  
 GN LIN-39 OR CEH-15 OR C07H6.7.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodermidae; Caenorhabditis.  
 NX NCBI\_TaxID=6239;

OK NCBI\_TaxID=6239;  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=93327430; PubMed=8101475;  
 RA Clark S.G., Chisholm A.D., Horvitz H.R.;

RT "Control of cell fates in the central body region of C. elegans by  
 the homeobox gene 11n-39."  
 RL Cell 74:43-55(1993).  
 RN [2]

RP SEQUENCE FROM N.A. PubMed=8101474;  
 RX MEDLINE=93327429; Wang B.B., Mueller-Immbergluck M.M., Austin J., Robinson N.T.,  
 Chisholm A.D., Kenyon C.,  
 RA "A homeotic gene cluster patterns the anteroposterior body axis of C.  
 elegans."  
 RT Cell 74:29-42(1993).  
 RL [3]

RP SEQUENCE FROM N.A. STRAIN=Bristol N2;  
 RC Macri C., Vaudin M.;  
 RA Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: CONTROLS THE MIGRATION OF NEUROBLASTS AND THE  
 SUBSEQUENT MID-BODY REGION-SPECIFIC DEVELOPMENT.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE ANTP HOMEOBOX FAMILY.  
 CC "DEFORMED" SUBFAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; L19639; AAC37168.1; -  
 DR EMBL; L19248; AAB04137.1; -  
 DR EMBL; AC006605; AAK85445.1; -  
 DR PIR; A40720; A40720.  
 DR HSP; P02833; 1SAN.  
 DR WormPep; C07H6.7; CE03975.  
 DR TRANSFAC; T03369; -  
 DR InterPro; IPR001827; Antennapedia.  
 DR InterPro; IPR001356; Homeobox.  
 DR Pfam; PF00046; homeobox; 1.  
 DR PRINTS; PR00025; ANTENNAPEDIA.  
 DR PRINTS; PR00024; HOMEOBOX.  
 DR PRODOM; PD000010; Homeobox; 1.  
 DR SMART; SM00389; HOK; 1.  
 DR PROSITE; PS00027; HOMEOBOX\_1; 1.  
 DR PROSITE; PS00071; HOMEOBOX\_2; 1.  
 DR PROSITE; PS00032; ANTENNAPEDIA; 1.  
 DR PROSITE; PS00032; ANTENNAPEDIA; 1.  
 KW Homeobox; DNA-binding; Nuclear protein; Developmental protein.  
 FT DOMAIN 20 36  
 FT DOMAIN 121 129  
 FT DOMAIN 148 153  
 FT DNA BIND 165 224  
 FT DOMAIN 228 248  
 FT SEQUENCE 253 AA; 27814 MW; 6B59DE9568B834A1 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 435 EGD DDDDD 441  
 Db 119 EGD DDDDD 125

RESULT 54  
 ID T227 RAT STANDARD; PRT; 253 AA.

AC 062880;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Transcription initiation factor TFIID 31 kDa subunit (TAFII-31)  
 DE (TAFII-32) (TAFII32) (Neuronal cell death related gene in neuron -7)

```

DE (DN-7).
GN TAF9 OR TAFI131.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RX SEQUENCE FROM N.A.
MEDLINE=97312553; PubMed=9168994;
RA Aoki T., Koike T., Nakano T., Shibahara K., Nishimura H., Kikuchi H.,
RA Honjo T.;
RT "Rat TAFI131 gene is induced upon programmed cell death in
RT differentiated PC12 cells deprived of NGF.";
RL Biochem. Biophys. Res. Commun. 234:230-234(1997).
CC -1- FUNCTION: TAFs ARE COMPONENTS OF THE TRANSCRIPTION FACTOR IID
CC (TFIID) COMPLEX THAT ARE ESSENTIAL FOR MEDIATING REGULATION OF RNA
CC POLYMERASE TRANSCRIPTION. TAFI131 IS A COACTIVATOR FOR THE P53
CC PROTEIN. ALSO INTERACTS WITH THE ACIDIC TRANSACTIVATOR VIRAL
CC PROTEIN 16 (VP16) AS WELL AS WITH THE GENERAL TRANSCRIPTION FACTOR
CC TFIIB (BY SIMILARITY).
CC -1- SUBUNIT: TFIID IS COMPOSED OF TATA BINDING PROTEIN (TBP) AND A
CC NUMBER OF TBP-ASSOCIATED FACTORS (TAFs) (BY SIMILARITY). TAFI131
CC AND P53 BIND TO EACH OTHER VIA AMINO ACID RESIDUES IN THE AMINO-
CC TERMINAL DOMAIN OF P53 THAT ARE ESSENTIAL FOR TRANSCRIPTION (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: BELONGS TO THE TAF2G FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U40188; AAC53201.1; -
DR InterPro; IPR004822; TFIID-31.
DR InterPro; IPR003162; TFIID-31.
DR Pfam; PF02291; TFIID-31.1.
DR ProDom; PD011023; TFIID-31; 1.
DR Transcription regulation; Nuclear protein.
KW DOMAIN 238 251 POLY-ASP.
FT SEQUENCE 253 AA; 27620 MW; 42045091A9B94378 CRC64;
SQ
Query Match 1.5%; Score 7; DB 1; Length 253;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
437 DDDDDDD 443
238 DDDDDDD 244

```

```

RESULT 55
ICP3_HSV1F STANDARD; PRT; 263 AA.
AC P08353;
DT 01-AUG-1998 (Rel. 08, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Infected cell protein ICP34.5 (Neurovirulence factor ICP34.5).
GN ICP34.5.
OS Herpes simplex virus (type 1 / strain F).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10304;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=6115412; PubMed=3003394;
RA Chou J., Roizman B.;
RT "The terminal a sequence of the herpes simplex virus genome contains
RT the promoter of a gene located in the repeat sequences of the L

```

```

RT component.
RL J. Virol. 57:629-637(1986).
RN [2]
RP REVISIONS.
RX MEDLINE=90156494; PubMed=2154589;
RA Chou J., Roizman B.;
RT "The herpes simplex virus 1 gene for ICP34.5, which maps in inverted
RT repeats, is conserved in several limited-passage isolates but not in
RT strain 17syn-".
RL J. Virol. 64:1014-1020(1990).
CC -1- FUNCTION: NOT KNOWN; HAS A PROFOUND EFFECT ON THE GROWTH OF THE
CC VIRUS IN CENTRAL NERVOUS SYSTEM TISSUE.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; IN INFECTED CELL.
CC -1- DEVELOPMENTAL STAGE: ACCUMULATES LATE IN INFECTION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M12240; AAA45794.1; ALT_SEQ.
DR EMBL; M33699; AAA45790.1; -
DR PIR; A27768; WMBE38.
KW Repeat; Late protein.
FT DOMAIN 161 190
SQ SEQUENCE 263 AA; 27533 MW; F5084106A08A8CB1 CRC64;
10 X 3 AA TANDEM REPEATS OF A-T-P.
QY 437 DDDDDDD 443
DB 75 DDDDDDD 81
Query Match 1.5%; Score 7; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
437 DDDDDDD 443
75 DDDDDDD 81

```

```

RESULT 56
OSTP_CHICK STANDARD; PRT; 264 AA.
AC P23498;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Osteopontin precursor (Bone sialoprotein I).
GN SPPI.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91236779; PubMed=2033080;
RA Castagnola P., Bet P., Quarto R., Gennari M., Migliaccio G.,
RA Canceda R.;
RT "cDNA cloning and gene expression of chicken osteopontin. Expression
RT of osteopontin mRNA in chondrocytes is enhanced by trypsin treatment
RT of cells."
RL J. Biol. Chem. 266:9944-9949(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91159433; PubMed=2001376;
RA Moore M.A., Gotoh Y., Rafidi K., Gerstenfeld L.C.;
RT "Characterization of a cDNA for chicken osteopontin: expression
RT during bone development, osteoblast differentiation, and tissue
RL Biochemistry 30:2501-2508(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=94192994; PubMed=8144023;

```





CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 CC Xenopodidae; Xenopus.  
 OK NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Oocyte;  
 RX MEDLINE=92375699; PubMed=1508694;  
 RA Morales J., Cormier P., Mulner-Lorillon O., Poulhe R.,  
 RT "Molecular cloning of a new guanine nucleotide-exchange protein, Efl  
 delta."  
 RN Nucleic Acids Res. 20:4091-4091(1992).  
 RL [2]  
 RP PRELIMINARY SEQUENCE OF 15-27; 112-118; 186-192 AND 226-259.  
 RX MEDLINE=9131196; PubMed=1869528;  
 RA Janssen G.M.C., Morales J., Schipper A., Labbes J.C.,  
 RA Mulner-Lorillon O., Belle R., Moeller W.,  
 RT "A major substrate of maturation promoting factor identified as  
 RT elongation factor 1 beta gamma delta in Xenopus laevis."  
 RL J. Biol. Chem. 266:14885-14888(1991).  
 CC -1- FUNCTION: EF-1-BETA AND EF-1-DELTA STIMULATE THE EXCHANGE OF  
 CC GDP BOUND TO EF-1-ALPHA TO GTP.  
 CC -1- SUBUNIT: EF-1 IS COMPOSED OF FOUR SUBUNITS: ALPHA, BETA,  
 CC DELTA, AND GAMMA.  
 CC -1- SIMILARITY: BELONGS TO THE EF-1-BETA/EF-1-DELTA FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X66837; CAA47313.1;  
 DR PIR; A39331; A39331.  
 DR PIR; S24635; S24635.  
 DR PIR; S26280; S26280.  
 DR InterPro; IPR001326; EF1-BD.  
 DR Pfam; PF00736; EF1BD\_1.  
 DR PROSITE; PS00824; EF1BD\_1; 1.  
 DR PROSITE; PS00825; EF1BD\_2; 1.  
 DR Elongation factor; protein biosynthesis.  
 KW SEQUENCE 265 AA; 29237 MW; 61A1898EC3F9E402 CRC64;  
 SQ  
 Query Match 1.5%; Score 7; DB 1; Length 265;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 437 DDDDDDD 443  
 |||||  
 Db 134 DDDDDDD 140  
 RESULT 59  
 HAX1 MOUSE STANDARD; PRT; 280 AA.  
 ID HAX1 MOUSE  
 AC O35387;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE HSI-associating protein X-1 (HAX-1) (HSI-binding protein).  
 GN HAX1.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OK NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Matanabe T., Takeshita H.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: DIRECTLY ASSOCIATES WITH HSI, THROUGH BINDING TO ITS N-  
 CC TERMINAL REGION. MAY FUNCTION IN PROMOTING CELL SURVIVAL. MAY ALSO  
 CC ASSOCIATE WITH CONTACTIN/EMSI IN NONLYMPHOID CELLS (BY

CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: MAINLY IN MITOCHONDRION, BUT ALSO TO A  
 CC LESSER EXTENT IN ENDOPLASMIC RETICULUM AND NUCLEAR ENVELOPE (BY  
 CC SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS.  
 CC -1- SIMILARITY: WEAK, TO THE NIP3 AND BCL-2 FAMILY PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF023482; AAB81081.1;  
 DR WGD; MGI:1346319; Hax1.  
 DR Mitochondrion.  
 FT DOMAIN 30 44 ASP/GLU-RICH (HIGHLY ACIDIC).  
 FT DOMAIN 115 280 INVOLVED IN HSI BINDING (BY SIMILARITY).  
 SQ SEQUENCE 280 AA; 31654 MW; 3123979BA10B45D1 CRC64;  
 Query Match 1.5%; Score 7; DB 1; Length 280;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 437 DDDDDDD 443  
 |||||  
 Db 30 DDDDDDD 36  
 RESULT 60  
 YNV9 CAEEL STANDARD; PRT; 289 AA.  
 ID YNV9 CAEEL  
 AC P34572; Q23407;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein T16H12.9 in chromosome III.  
 GN T16H12.9  
 OS Caenorhabditis elegans.  
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 CC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OK NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Smith A., Berke W.;  
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP REVISIONS.  
 RA Durbin R.;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z30662; CAA83142.2;  
 DR EMBL; Z47357; CAA83142.2; JOINED.  
 DR EMBL; Z47357; CAA87428.2;  
 DR EMBL; Z30662; CAA87428.2; JOINED.  
 DR PIR; S42381; S42381.  
 DR Wormpep; T16H12.9; CE28843.  
 DR Hypothetical protein.  
 KW DOMAIN 134 145 ASP-RICH.  
 FT DOMAIN 134 145  
 SQ SEQUENCE 289 AA; 32186 MW; 425B5C6107108D7F CRC64;  
 Query Match 1.5%; Score 7; DB 1; Length 289;

Best Local Similarity 100.0%; Pred.No. 41;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
Db 134 DDDDDDD 140

RESULT 61

UBC3\_YEAST STANDARD; PRT; 295 AA.  
ID UBC3\_YEAST  
AC P14682;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Ubiquitin-conjugating enzyme E2-34 kDa (EC 6.3.2.19)  
DE (ubiquitin-conjugating enzyme E2-34 kDa (EC 6.3.2.19)  
DE control protein 34).  
GN UBC3 OR CDC34 OR DNA6 OR YDR054C OR YD9609.08C OR D4211.  
OS Saccharomyces cerevisiae (Baker's yeast).  
SS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
NCBI\_TaxID=4932;  
[1]  
SEQUENCE FROM N.A.  
RX MEDLINE=86321694; PubMed=2842867;  
RA Goshl M.G., Yochem J., Mergath J.P., Varshavsky A.,  
Byers B.;  
RT "The yeast cell cycle gene CDC34 encodes a ubiquitin-conjugating  
enzyme.";  
RL Science 241:1331-1335 (1988).  
[2]  
SEQUENCE FROM N.A.  
RX MEDLINE=96381250; PubMed=8789263;  
RA Brandt P., Ramlow S., Otto B., Bloeker H.;  
RT "Nucleotide sequence analysis of a 32,500 bp region of the right arm  
of Saccharomyces cerevisiae chromosome IV.";  
RL Yeast 12:85-90 (1996).  
[3]  
SEQUENCE FROM N.A.  
RC STRAIN=5288C / AB972;  
RA Hunt S., Bowman S., Harris D., Barrell B.G., Rajandream M.A.;  
RL Submitted (May-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO  
OTHER PROTEINS. CAPABLE, IN VITRO, TO UBIQUITINATE HISTONE H2A.  
CC -1- FUNCTION: MEDIATES THE INITIATION OF DNA REPLICATION (TRANSITION  
OF G1 TO S PHASE IN CELL CYCLE).  
CC -1- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +  
diphosphate + protein N-ubiquityllysine.  
CC -1- PATHWAY: Ubiquitin conjugation; second step.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- DOMAIN: THE ACIDIC C-TERMINAL EXTENSION IS ESSENTIAL FOR THE CELL  
CYCLE FUNCTION.  
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR  
UBIQUITIN-THIOLESTER FORMATION.  
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
EMBL: M21877; AAA35188.1; -  
EMBL: X84162; CAA58970.1; -  
EMBL: Z74350; CAA98872.1; -  
EMBL: Z49209; CAA90083.1; -  
EMBL: A41241; A41241.  
DR PIR; A41241; A41241.  
DR HSSP; 002159; 2UCZ.  
DR SGD; S0002461; CDC34.  
DR InterPro; IPR000608; UBC3\_conjugat.

DR Pfam; PF00179; UO\_conj\_1.  
DR ProDom; PD000461; UBC3\_conjugat; 1.  
DR SMART; SM00212; UBC3; 1.  
DR PROSITE; PS00183; UBIQUITIN\_CONJUGAT\_1; 1.  
DR PROSITE; PS0127; UBIQUITIN\_CONJUGAT\_2; 1.  
KW Ubl conjugation pathway; ligase; DNA replication; Nuclear protein;  
KW Cell cycle; Cell division; Multigene family.  
FT BINDING 95 95 UBIQUITIN (BY SIMILARITY).  
FT DOMAIN 121 289 ASP/GLU-RICH (ACIDIC).  
FT SEQUENCE 295 AA; 34064 MW; 1CE3E0C3AB1436DC CRC64;  
SQ

Query Match 1.5%; Score 7; DB 1; Length 295;  
Best Local Similarity 100.0%; Pred.No. 42;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
Db 255 DDDDDDD 261

RESULT 62

PEX5\_CANAL STANDARD; PRT; 296 AA.  
ID PEX5\_CANAL  
AC 074711;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Peroxisomal targeting signal receptor (Peroxisomal protein PAS10)  
DE (peroxin-5) (PTS1 receptor) (Fragment).  
GN PEX5.  
OS Candida albicans (Yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
NCBI\_TaxID=5476;  
[1]  
SEQUENCE FROM N.A.  
RC STRAIN=ATCC 64385 / 1001;  
RX MEDLINE=98451822; PubMed=9778800;  
RA Navarro-Garcia F., Perez-Diaz R., Negro A., Pla J., Nombela C.;  
RT "Cloning and sequence of a 3.835 kbp DNA fragment containing the HIS4  
gene and a fragment of a PEX5-like gene from Candida albicans.";  
RL Yeast 14:1147-1157 (1998).  
CC -1- FUNCTION: BINDS TO THE C-TERMINAL PTS1-TYPE TRIPEPTIDE PEROXISOMAL  
TARGETING SIGNAL (SKL-TYPE) AND PLAYS AN ESSENTIAL ROLE IN  
PEROXISOMAL PROTEIN IMPORT (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: ITS DISTRIBUTION APPEARS TO BE DYNAMIC. IT  
IS PROBABLY A CYCLING RECEPTOR FOUND MAINLY IN THE CYTOPLASM AND  
AS WELL ASSOCIATED TO THE PEROXISOMAL MEMBRANE THROUGH A DOCKING  
FACTOR (PEX13) (BY SIMILARITY).  
CC -1- SIMILARITY: CONTAINS 8 TPR REPEATS.  
CC -1- SIMILARITY: STRONG, TO OTHER PEROXISOMAL TARGETING SIGNAL  
REPEATS.  
CC -----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
EMBL: AJ003115; CAA05870.1; -  
DR InterPro; IPR001440; TPR.  
DR Pfam; PF00515; TPR; 4.  
DR SMART; SM00028; TPR; 4.  
KW Peroxisome; Repeat; TPR repeat; Transport; Protein transport.  
FT NON TER 1 1  
FT REPEAT 1 33 TPR 2.  
FT REPEAT 34 67 TPR 3.  
FT REPEAT 68 105 TPR 4.  
FT REPEAT 106 143 TPR 5.  
FT REPEAT 144 177 TPR 6.  
FT REPEAT 178 211 TPR 7.

FT REPEAT 212 245 TPR 8.  
 SQ SEQUENCE 296 AA; 33220 MW; 1F78FAE76395A0B CRC64;  
 Query Match 1.5%; Score 7; DB 1; Length 296;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 95 NSNRSEE 101  
 Db 190 NSNRSEE 196

RESULT 63  
 MURB\_BACSU STANDARD; PRT; 303 AA.  
 ID MURB\_BACSU  
 AC P18579; P16669; P37581;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 UDP-N-acetylglucosaminylglucosamine reductase (EC 1.1.1.158) (UDP-N-acetylglucosamine dehydrogenase).  
 MURB

OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90078133; PubMed=2556375;  
 RA Beal B., Lutkenhaus J.;  
 RT "Nucleotide sequence and insertional inactivation of a Bacillus subtilis gene that affects cell division, sporulation, and temperature sensitivity";  
 RL J. Bacteriol. 171:6821-6834(1989).  
 RN [2]  
 RP SEQUENCE OF 1-23 FROM N.A.  
 RX MEDLINE=92380484; PubMed=1387377;  
 RA Miyao A., Yoshimura A., Sato T., Yamamoto T., Theeragool G., Kobayashi Y.;  
 RT "Sequence of the Bacillus subtilis homolog of the Escherichia coli cell-division gene *murG*";  
 RL Gene 118:147-148(1992).  
 RN [3]  
 RP SEQUENCE OF 250-303 FROM N.A.  
 RX MEDLINE=90078134; PubMed=2556376;  
 RA Harry E.J., Wake R.G.;  
 RT "Cloning and expression of a Bacillus subtilis division initiation gene for which a homolog has not been identified in another organism";  
 RL J. Bacteriol. 171:6835-6839(1989).  
 RN [4]  
 RP FUNCTION.  
 RX MEDLINE=96060848; PubMed=7590298;  
 RA Rowland S.L., Errington J., Wake R.G.;  
 RT "The Bacillus subtilis cell-division 135-137 degrees region contains an essential *orf* with significant similarity to *murB* and a dispensable *sdp* gene";  
 RL Gene 164:113-116(1995).  
 CC -1- FUNCTION: CELL WALL FORMATION.  
 CC -1- CATALYTIC ACTIVITY: UDP-N-acetylglucosamine + NADP(+) = UDP-N-acetyl-3-O-(1-carboxyvinyl)-D-glucosamine + NADPH.  
 CC -1- COFACTOR: P4D.  
 CC -1- PATHWAY: Peptidoglycan biosynthesis.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (probable).  
 CC -1- SIMILARITY: BELONGS TO THE MURB FAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M31827; AAA83969.1; -  
 DR EMBL; D10602; BAA01455.1; -  
 DR EMBL; M31800; AAA22392.1; -  
 DR EMBL; 299111; CAB13396.1; -  
 DR PIR; S26500; S26500.  
 DR PIR; A43727; A43727.  
 DR PIR; PC1128; PC1128.  
 DR Subtilast; BG10228; murB.  
 DR InterPro; IPR003170; murB.  
 DR InterPro; IPR001575; Oxid\_FAD bind.  
 DR Pfam; PF01565; FAD binding\_4; 1.  
 DR Pfam; PF02873; MurB\_C; 1.  
 DR TIGRFAMs; TIGR00179; murB; 1.  
 DR Peptidoglycan synthetase; Cell wall; Cell division; Oxidoreductase;  
 KW NADP, Flavoprotein; P4D; Complete proteome.  
 SQ SEQUENCE 303 AA; 32808 MW; C3F5367C84E8A6E CRC64;  
 Query Match 1.5%; Score 7; DB 1; Length 303;  
 Best Local Similarity 100.0%; Pred. No. 43;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 306 QELKERE 312  
 Db 6 QELKERE 12

RESULT 64  
 CDK3\_HUMAN STANDARD; PRT; 305 AA.  
 ID CDK3\_HUMAN  
 AC Q00526;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Cell division protein kinase 3 (EC 2.7.1.-).  
 GN CDK3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal brain;  
 RX MEDLINE=92347325; PubMed=1639063;  
 RA Meyerson M., Enders G.H., Wu C.-L., Su L.-K., Gotta C., Nelson C., Harlow E., Tsai L.-H.;  
 RT "A family of human cdc2-related protein kinases";  
 RL EMBO J. 11:2909-2917(1992).  
 CC -1- FUNCTION: PROBABLY INVOLVED IN THE CONTROL OF THE CELL CYCLE.  
 CC INTERACTS WITH A YET UNKNOWN TYPE OF CYCLIN. CAN PHOSPHORYLATE HISTONE H1.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC CDC2/CDK3 SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X66357; CAA47001.1; -  
 DR PIR; S22743; S22743.  
 DR PIR; S23382; S23382.  
 DR HSSP; P24941; ICKP.  
 DR Genew; HGNC:1772; CDK3.  
 DR MIM; 123828; -  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SMO0220; S\_TKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.

DR PROSITE; PS00108; PROTEIN KINASE ST. 1.  
 DR PROSITE; PS00011; PROTEIN KINASE\_DOM; 1.  
 KM Transferrin; Serine/threonine-protein kinase; ATP-binding;  
 KM Cell cycle; Cell division; Mitosis; Phosphorylation.  
 FT DOMAIN 4 286  
 FT NP BIND 10 18  
 FT BINDING 33 33  
 FT ACT\_SITE 127 127  
 FT MOD\_RES 14 14  
 FT MOD\_RES 15 15  
 FT MOD\_RES 160 160  
 FT MOD\_RES 35045 AA; 1128BE0096A262A CRC64;  
 SQ SEQUENCE 305 AA; 35045 MW; 1128BE0096A262A CRC64;  
 Query Match 1.5%; Score 7; DB 1; Length 305;  
 Best Local Similarity 100.0%; Pred. No. 43;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 129 VHRDLK 135  
 123 VHRDLK 129  
 RESULT 65  
 MLF\_DROME STANDARD; PRT; 309 AA.  
 ID OGNKVO; Q9V7G3;  
 AC 15-UN-2002 (Rel. 41, Last sequence update)  
 DT 15-UN-2002 (Rel. 41, Last sequence update)  
 DT 15-UN-2002 (Rel. 41, Last sequence update)  
 DE Myeloid leukemia factor (Myelodysplasia-myeloid leukemia factor) (DMF).  
 GN MLF OR CG8295.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 CC NCBI\_TaxID=7227;  
 [1]  
 SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RP MEDLINE=20578696; PubMed=1137299;  
 RX Ono K., Takahashi Y., Hirose F., Inoue Y.H., Taguchi O., Nishida Y.,  
 RA Matsukage A., Yamaguchi M.,  
 RT "Characterization of a Drosophila homologue of the human  
 myelodysplasia/myeloid leukemia factor (MLF).";  
 RL Gene 260:133-143(2000).  
 [2]  
 SEQUENCE FROM N.A.  
 RP STRAIN=Berkley;  
 AC MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers J.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Aitell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bhotishvili S.,  
 RA Borkova D., Botchan M.R., Bouck B.P., Brokstein P., Brotler P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Chert J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Maye A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fodor C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris W.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houson K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalaiah M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laake P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshirel A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuske D.R., Pacle J.M.,  
 RA Palazkzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier B., Spalding A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Morley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 [3]  
 RP REVISIONS.  
 RC STRAIN=Berkley;  
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C.M., Bernan B.P., Carlson J.W., Celisner S.E.,  
 RA Clamp M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,  
 RA Harris N.L., Krommiller B., Marshall B., Milburn G.H., Richter J.,  
 RA Russo S., Seale S.M.J., Smith E., Shu S., Smutnak F.,  
 RA Whitfield E.J., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Mungall C.J., Lewis S.E.;  
 RT "Annotation of Drosophila melanogaster genome.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Interacts with DRB-binding  
 CC factor (DRBF).  
 CC -1- TISSUE SPECIFICITY: Expressed at high levels in unfertilized eggs,  
 CC early embryos, pupae and adult males while a low level expression  
 CC is found in adult females and larvae.  
 CC -1- DEVELOPMENTAL STAGE: High levels are seen in unfertilized eggs and  
 CC expression increases slightly during early embryo stages (2-3  
 CC hrs). Levels are high in embryos until 4 hrs after fertilization  
 CC and then decrease gradually through embryonic and larval stages.  
 CC -1- SIMILARITY: BELONGS TO THE MLF FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; AB043986; BAA6391.1; -;  
 DR EMBL; AB003809; AAF58093.2; -;  
 DR FLYBASE; FBgn0034051; MLF.  
 FT DOMAIN 96 202 INTERACTION WITH DREF.  
 FT MOD\_RES 236 243 POLY-ASP  
 FT MOD\_RES 34413 MW; 4C50951D548FD9AD CRC64;  
 SQ SEQUENCE 309 AA; 34413 MW; 4C50951D548FD9AD CRC64;  
 Query Match 1.5%; Score 7; DB 1; Length 309;  
 Best Local Similarity 100.0%; Pred. No. 44;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 437 DDDDDDD 443  
 236 DDDDDDD 242  
 RESULT 66  
 DOS2 YEAST STANDARD; PRT; 310 AA.  
 ID DOS2 YEAST  
 AC P54858;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE DOS2 protein.  
 GN DOS2 OR YDR068W OR YD9609.22 OR YD8554.01 OR D4267.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

CC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxId=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Singer J.D., Manning B.E., Formosa T.;  
 RT "Control of single-copy DNA replication requires genes that act in  
 RL ubiquitin metabolism."; to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=96381250; PubMed=8769263;  
 RX Brandt P., Ramlow S., Otto B., Bloecher H.;  
 RT "Nucleotide sequence analysis of a 32,500 bp region of the right arm  
 RL of Saccharomyces cerevisiae chromosome IV."; Yeast 12:85-90(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=5288c / AB972;  
 RX Hunt S., Bowman S., Harris D., Barrell B.G., Rajandream M.A.;  
 RT Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: ACTS IN UBIQUITIN METABOLISM AND IS NECESSARY FOR THE  
 CC CONTROL OF SINGLE-COPY DNA REPLICATION.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL, U9857; AAA66522.1; -  
 DR EMBL, X84162; CA58984.1; -  
 DR EMBL, 274364; CA58984.1; -  
 DR EMBL, 249209; CA48909.1; -  
 DR EMBL, 246796; CA486790.1; -  
 DR SGD, S0002475; DOS2.  
 FT DOMAIN 255 259 POLY-GLY.  
 FT DOMAIN 302 308 POLY-ASP.  
 SQ SEQUENCE 310 AA; 35964 MW; B9B318093E984713 CRC64;  
 Query Match 1.5%; Score 7; DB 1; Length 310;  
 Best Local Similarity 100.0%; Pred. No. 44;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 437 DDDDDDD 443  
 302 DDDDDD 308  
 RESULT 67  
 MK03 MOUSE STANDARD; PRT; 314 AA.  
 ID MK03 MOUSE  
 AC 063844; 061531;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Mitogen-activated protein kinase 3 (EC 2.7.1.1.) (Extracellular signal-  
 DE regulated kinase 1) (ERK-1) (Insulin-stimulated MAP2 kinase) (MAP  
 DE kinase 1) (MAPK 1) (P44-ERK1) (ERK2) (P44-MAPK) (Microtubule-  
 DE associated protein-2 kinase) (MNK1) (Fragments).  
 GN MAPK3 OR PRK3 OR ERK1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE OF 1-10 FROM N.A.  
 RC TISSUE=Pre-B cell;  
 RX MEDLINE=92020947; PubMed=1717989;  
 RA Crewe C.M., Alessandrini A.A., Erikson R.L.;  
 RT "Mouse Erk-1 gene product is a serine/threonine protein kinase that  
 has the potential to phosphorylate tyrosine.";

RL Proc. Natl. Acad. Sci. U.S.A. 88:8845-8849(1991).  
 RN [2]  
 RP SEQUENCE OF 11-314 FROM N.A.  
 RC TISSUE=Fetal brain;  
 RX MEDLINE=91369479; PubMed=1716439;  
 RA de Miguel C., Klisman D., Patel J., Deterra-Wadleigh S.D.;  
 RT "Molecular analysis of microtubule-associated protein-2 kinase cDNA  
 RL from mouse and rat brain."; DNA Cell Biol. 10:505-514(1991).  
 RN [3]  
 RP SEQUENCE OF 105-143 FROM N.A.  
 RC STRAIN=CBA; TISSUE=Bone marrow;  
 RX MEDLINE=93185941; PubMed=8444355;  
 RA Erhler M.A., Nagorskaya T.V., Visser J.W.M., Belyavsky A.V.;  
 RT "Novel CDC2-related protein kinases produced in murine hematopoietic  
 RL stem cells."; Gene 124:305-306(1993).  
 RN [4]  
 RP SEQUENCE OF 105-139 FROM N.A.  
 RX MEDLINE=93092802; PubMed=1459009;  
 RA Erhler M.A., Nagorskaya T.V., Visser J.W.M., Belyavsky A.V.;  
 RT "Identification of new protein kinase genes, similar to kinases of  
 RL the cdc2 family and expressed in murine hematopoietic stem cells.";  
 CC Dokl. Akad. Nauk SSSR 324:893-897(1992).  
 CC -1- FUNCTION: PHOSPHORYLATES MICROTUBULE-ASSOCIATED PROTEIN-2 (MAP2).  
 CC MYELIN BASIC PROTEIN (MBP), AND ERK-1; MAY PROMOTE ENTRY IN THE  
 CC CELL CYCLE.  
 CC -1- ENZYME REGULATION: ACTIVATED BY TYROSINE AND THREONINE  
 CC PHOSPHORYLATION.  
 CC -1- PTM: AUTOPHOSPHORYLATED ON THREONINE AND TYROSINE RESIDUES.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MAP KINASE SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL, S58470; AAB1973.1; -  
 DR EMBL, X64605; CA45889.1; -  
 DR HSSP, P27703; 2ERK.  
 DR MGD, MGI:1346859; MAPK3.  
 DR InterPro: IPR000719; MAPK3.  
 DR InterPro: IPR003527; MAP\_Kin.  
 DR InterPro: IPR002290; Ser\_thr\_kinase.  
 DR ProDom: PD000001; Euk\_pkinase; 1.  
 DR PROSITE, PS01351; MAPK; 1.  
 DR PROSITE, PS00107; PROTEIN KINASE ATP, PARTIAL.  
 DR PROSITE, PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE, PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding; Cell cycle;  
 KW Phosphorylation.  
 FT NON\_TER 1 1  
 FT DOMAIN 1 4 POLY-GLY.  
 FT NON\_CONS 10 11  
 FT MOD\_RES 137 137 PHOSPHORYLATION (ACTIVATES THE KINASE)  
 FT MOD\_RES 137 137 PHOSPHORYLATION (ACTIVATES THE KINASE)  
 FT MOD\_RES 139 139 PHOSPHORYLATION (ACTIVATES THE KINASE)  
 FT MOD\_RES 139 139 PHOSPHORYLATION (ACTIVATES THE KINASE)  
 FT CONFLICT 112 112 T -> P (IN REF. 3 AND 4).  
 FT SEQUENCE 314 AA; 36198 MW; F06085E047BF7C22 CRC64;  
 Query Match 1.5%; Score 7; DB 1; Length 314;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 147 LKICDFG 153  
 115 LKICDFG 121



OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxId=158879, 196620, 1280;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=N315;  
RA MEDLINE=21311952; PubMed=11418146;  
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,  
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshiyama A.,  
RA Mizutani H., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
RA Sekimizu K., Hiraoka H., Kohara S., Goto S., Yabuzaki J.,  
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
RA Hattori M., Ogasawara N., Hayashi H., Hiratake K.;  
RA "Whole genome sequencing of methicillin-resistant Staphylococcus  
RT aureus".  
RL Lancet 357:1225-1240 (2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MM2;  
RA MEDLINE=22040717; PubMed=12044378;  
RA Baba T., Takuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,  
RA Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,  
RA Yamamoto K., Hiratake K.;  
RA "Genome and virulence determinants of high virulence community-  
RT acquired MRSA".  
RL Lancet 359:1819-1827 (2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=93138806; PubMed=8423103;  
RA Cooney J.C., Klenke Z., Foster T.J., O'Toole P.W.;  
RT "The gamma-hemolysin locus of Staphylococcus aureus comprises three  
RT linked genes, two of which are identical to the genes for the F and S  
RL components of leukocidin".  
RN Infect. Immun. 61:768-771 (1993).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=89310396; PubMed=3075655;  
RA Cooney J.C., Mulvey M., Arbuthnot J., Foster T.;  
RT "Molecular cloning and genetic analysis of the determinant for gamma-  
RT lyisin, a two-component toxin of Staphylococcus aureus".  
RL J. Gen. Microbiol. 134:2179-2188 (1988).  
CC - FUNCTION: GAMMA-HEMOLYSIN CAUSES HEMOLYSIS IN RED BLOOD CELLS.  
CC - SUBUNIT: CONSISTS OF TWO PROTEIN COMPONENTS (H-GAMMA-I AND -II).  
CC H-GAMMA-I IS ALSO THE F SUBUNIT OF LEUKOCIDIN.  
CC - SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as their content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AP003137; BAB93511.1;  
DR EMBL; AP004830; BAB96209.1;  
DR EMBL; L01055; AAA26639.1;  
DR HSSP; P09616; 7AHL.  
DR InterPro; IPR001776; Aerolysin.  
DR InterPro; IPR001340; Hemolysin\_pore.  
DR InterPro; IPR003963; Staph\_blon\_tcn.  
DR Pfam; PF01171; Aerolysin\_1.  
DR PRINTS; PR01468; BICOMPNTOXIN.  
DR TIGRFAMs; TIGR01002; hlyII, 1.  
DR Hemolysin; Toxin; Signal; Complete proteome.  
FT SIGNAL 1 25  
FT CHAIN 26 325 GAMMA-HEMOLYSIN COMPONENT B.  
FT SEQUENCE 325 AA; 36711 MM; 082999EB9339B6A0 CRC64;  
Query Match 1.54; Score 7; DB 1; Length 325; Pred. No. 46;  
Best Local Similarity 100.0%; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 LSVLSHR 63  
 |||||  
 Db 252 LSVLSHR 258

RESULT 71  
 KRAB\_MOUSE  
 ID KRAB\_MOUSE STANDARD; PRT; 328 AA.  
 AC P28028;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE B-Raf proto-oncogene serine/threonine-protein kinase (EC 2.7.1.-)  
 DE (Fragment).  
 DE BRAF OR B-RAF.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=91271351; PubMed=2052597;  
 RX Miki T., Fleming T.P., Crescenzi M., Molloy C.J., Blam S.B.,  
 RA Reynolds S.H., Aaronson S.A.;  
 RT "Development of a highly efficient expression cDNA cloning system:  
 RT application to oncogene isolation.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:5167-5171 (1991).  
 CC -1- FUNCTION: INVOLVED IN THE TRANSDUCTION OF MITOGENIC SIGNALS  
 CC FROM THE CELL MEMBRANE TO THE NUCLEUS.  
 CC -1- FUNCTION: MAY PLAY A ROLE IN THE POSTSYNAPTIC RESPONSES OF  
 CC HIPPOCAMPAL NEURON.  
 CC -1- DISEASE: PARTICIPATES IN A CHROMOSOMAL TRANSLOCATION THAT PRODUCES  
 CC A T(11A-BRAF (718) ONCOGENE ORIGINALLY ISOLATED FROM A FIBRILAR-  
 CC INDUCED HEPATOMA.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MIL/RAF SUBFAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by, and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>).  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----

EMBL: M64429; AAA37320.1; ALT\_INIT.  
 DR PIR; A40951; TVMSBF.  
 DR HSSP; P08631; IAD5.  
 DR MGD; MG1:88190; BraF.  
 DR InterPro; IPR000719; Euk\_Pkinase.  
 DR InterPro; IPR004040; STY\_Pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_Pkinase.  
 DR Pfam; PF00069; Pkinase; 1.  
 DR ProDom; PDO00001; Euk\_Pkinase; 1.  
 DR SMART; SMO0221; STYK; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Transferase; Serine/threonine-protein kinase; Proto-oncogene;  
 KW ATP-binding; Chromosomal translocation.  
 FT NON\_TER 1  
 FT DOMAIN 19 279 PROTEIN\_KINASE.  
 FT NP\_BIND 25 33 ATP (BY SIMILARITY).  
 FT BINDING 45 45 ATP (BY SIMILARITY).  
 FT ACT\_SITE 138 138 BY SIMILARITY.  
 SQ SEQUENCE 328 AA; 36986 MW; 67A2BFBF78A78B3D CRC64;

Query Match 1.5%; Score 7; DB 1; Length 328;  
 Best Local Similarity 100.0%; Pred. No. 46;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 IHRDPS 136

Db 135 IHRDLKS 141

RESULT 72  
AAIP WHEAT STANDARD; PRT; 332 AA.  
AC 002066;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Abiotic acid-inducible protein kinase (EC 2.7.1.-) (Fragment).  
OS Triticum aestivum (Wheat).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Triticeae; Triticum.  
OX NCB1\_TaxID=4565;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=cv. Brevor; TISSUE=seed;  
RC MEDLINE=9306204; PubMed=1438207;  
RX Anderson R.J., Walker-Simmons M.K.;  
RT "Isolation of a wheat cDNA clone for an abiotic acid-inducible  
transcript with homology to protein kinases."  
Proc. Natl. Acad. Sci. U.S.A. 89:10183-10187(1992).  
CC -1- FUNCTION: IS INVOLVED IN WATER-STRESS RESPONSES.  
CC -1- INDUCTION: BY ABSCISIC ACID AND DEHYDRATION.  
CC -1- PTM: AUTOPHOSPHORYLATED (POTENTIAL).  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: M94726; AAA96325.1; -  
DR PIR: A46408; A46408.  
DR HSSP: Q83450; 1A06.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR002250; Ser\_thr\_pkinase.  
DR Pfam: PF00069; pkinase; 1.  
DR ProDom: PD000001; Euk\_pkinase; 1.  
DR SMART: SM00220; S\_TKc; 1.  
DR PROSITE: PS00107; PROTEIN KINASE ATP; PARTIAL.  
DR PROSITE: PS00108; PROTEIN KINASE ST; 1.  
DR PROSITE: PS50011; PROTEIN KINASE DOM; 1.  
DR Transferrase; Serine/threonine-protein kinase; ATP-binding;  
phosphorylation.  
CC  
CC NON TER 1  
CC FT DOMAIN <1 250 PROTEIN KINASE.  
CC FT NP BIND 8 ATP (BY SIMILARITY).  
CC FT BINDING 23 ATP (BY SIMILARITY).  
CC FT ACT\_SITE 113 BY SIMILARITY.  
CC FT ACT\_SITE 113  
CC SQ SEQUENCE 332 AA; 37516 MW; E19ACAEIC0F0C837 CRC64;  
Query Match 1.5%; Score 7; DB 1; Length 332;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 147 LKICDPG 153  
Db 129 LKICDPG 135

DE Cell division control protein 2 homolog (EC 2.7.1.-) (p34 protein  
kinase).  
GN NCC-1 OR T05G5.3.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCB1\_TaxID=6239;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=Bristol N2;  
RC MEDLINE=94150718; PubMed=7906398;  
RX Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
Cratton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,  
Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
Johnston L., Jones M., Kershaw P., Kirsten J., Laister N.,  
Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,  
Sims M., Smailon N., Smith A., Smith M., Sonhammer E., Staden R.,  
Sulton J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,  
Wohlman P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans."  
Nucleic Acids Res. 21:3683-3692(1993).  
[2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=Bristol N2;  
RC Ferriz C., Thierry-Mieg D., le Peuch C.J.;  
RX Submitted (SEP-1992) to the EMBL/Genbank/DBJ databases.  
[3]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=95131956; PubMed=7830726;  
RX Mori H., Palmer R.E., Sternberg P.W.;  
RT "The identification of a Caenorhabditis elegans homolog of p34cdc2  
kinase."  
Mol. Gen. Genet. 245:781-786(1994).  
[4]  
RN SEQUENCE FROM N.A.  
RP STRAIN=Bristol N2;  
RC MEDLINE=99225469; PubMed=10207147;  
RX Boxem M., Srinivasan D.G., van den Heuvel S.;  
RT "The Caenorhabditis elegans gene ncc-1 encodes a cdc2-related kinase  
required for M phase in meiotic and mitotic cell divisions, but not  
for S phase."  
Development 126:2227-2239(1999).  
CC -1- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC  
CELL CYCLE. IT IS REQUIRED IN HIGHER CELLS FOR ENTRY INTO S-PHASE  
AND MITOSIS. P34 IS A COMPONENT OF THE KINASE COMPLEX THAT  
PHOSPHORYLATES THE REPLICATIVE CARBOXYL-TERMINUS OF RNA  
POLYMERASE II (BY SIMILARITY)  
CC -1- ENZYME REGULATION: PHOSPHORYLATION INACTIVATES THE ENZYME  
(BY SIMILARITY).  
CC -1- SUBUNIT: FORMS A STABLE BUT NON-COVALENT COMPLEX WITH A REGULATORY  
SUBUNIT AND WITH A CYCLIN.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC CDC2/CDCX SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: Z27079; CAA81590.1; -  
DR EMBL: X68384; CAA48455.1; -  
DR EMBL: S75262; AAC60520.1; -  
DR EMBL: AF129109; AAD37119.1; -  
DR PIR: S26572; S26572.  
DR PIR: S41003; S41003.  
DR HSSP: P24941; 1CKP.



DR WormPep; T0565.3; CE00315.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF00069; Pkinase.1.  
 DR ProDom; PD000001; Euk\_pkinase.1.  
 DR SMART; SMO0230; S\_TKC.1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP.1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST.1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM.1.  
 DR Transferrase; Serine/threonine-protein kinase; ATP-binding;  
 Cell cycle; Cell division; Mitosis; Phosphorylation.  
 FT DOMAIN 22 312  
 FT NP\_BIND 28 36  
 FT BINDING 31 51  
 FT ACT\_SITE 146 146  
 FT MOD\_RES 32 32  
 FT MOD\_RES 33 33  
 FT CONFLICT 176 176  
 SQ SEQUENCE 332 AA; 38295 MW; B6237E92949C8206 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 332;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 129 VIHRLK 135  
 Db 142 VIHRLK 148

RESULT 74  
 Y713\_CHLMU STANDARD; PRT; 332 AA.  
 ID Y713\_CHLMU  
 AC 09P0J3;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein TC0713.  
 GN TC0713.  
 OS Chlamydia muridarum.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_Taxid=83560;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MOPn / N1G9;  
 RX MEDLINE=20150255; PubMed=10684935;  
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidberg J.F.,  
 White O., Hickey E.K., Peterson J., Uteerback T., Berty K., Baas S.,  
 Liner K., Weidman J., Knout H., Craven B., Bowman C., Dodson R.,  
 Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,  
 Eilen J., Fraser C.M.;  
 "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia  
 pneumoniae AR39.";  
 RT Nucleic Acids Res. 28:1397-1406(2000).  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE UPF0158 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).  
 CC -----  
 CC EMBL; AE002339; AAF39526.1; -  
 DR TIGR; TC0713; -  
 DR InterPro; IPR005361; UPF0158.  
 DR Pfam; PF03682; UPF0158.1.  
 KW Hypothetical protein; Complete proteome.  
 FT DOMAIN 301 316  
 FT DOMAIN 308 316  
 SQ SEQUENCE 332 AA; 39202 MW; 27527DC5E6ABDEA4 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 332;

Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDD 443  
 Db 308 DDDDDDD 314

RESULT 75  
 SPT2\_YEAST STANDARD; PRT; 333 AA.  
 ID SPT2\_YEAST  
 AC P06843;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE SPT2 protein (Negative regulator of ty. transcription).  
 GN SPT2 OR SPM2 OR SIN1 OR YER161C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_Taxid=4932;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2991744; PubMed=2991744;  
 RA Roeder G.S., Beard C., Smith M., Keranen S.;  
 RT "Isolation and characterization of the SPT2 gene, a negative  
 RT regulator of Ty-controlled yeast gene expression.";  
 RL Mol. Cell. Biol. 5:1543-1553(1985).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91304406; PubMed=2072912;  
 RA Kruger W., Herkowitz I.;  
 RT "A negative regulator of HO transcription, SIN1 (SPT2), is a  
 RT nonspecific DNA-binding protein related to HMGI.";  
 RL Mol. Cell. Biol. 11:4135-4146(1991).  
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RA Mulligan J.T., Dietrich F.S., Hennessey K.M., Sehl P., Komp C.,  
 RA Wei Y., Taylor P., Nakahara K., Roberts D., Davis R.W.;  
 RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.  
 RN (4)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c / AB972;  
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,  
 RA Aviles E., Berro A., Brennan T., Carpenter J., Chen E., Cherry J.M.,  
 RA Chung E., Duncan M., Guman E., Hartzell G., Hunnicke-Smith S.,  
 RA Hymen R., Kayser A., Komp C., Lastkari D., Lew H., Lin D.,  
 RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,  
 RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,  
 RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: CHROMATIN PROTEIN. GLOBAL REGULATORY PROTEIN THAT PLAYS  
 CC POSITIVE AS WELL AS NEGATIVE REGULATORY ROLES IN TRANSCRIPTION.  
 CC -1- SUBUNIT: INTERACTS WITH SAPI AND CDC23.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -----  
 CC -1- SIMILARITY: PARTIAL (DNA-BINDING MOTIF) TO HMG PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).  
 CC -----  
 CC EMBL; M1165; AAA5083.1; -  
 DR EMBL; U18917; AAB64688.1; -  
 DR PIR; A23438; A23438.  
 DR PIR; S30813; S30813.  
 DR SGD; S0000963; SPT2.  
 KW DNA-binding; Activator; Repressor; Transcription regulation;  
 KW Nuclear protein.  
 SQ SEQUENCE 333 AA; 38551 MW; 60C3BED7E9A9535 CRC64;

Sun May 4 10:01:08 2003

us-09-757-982-5.oligo.rsp

Page 40

Query Match 1.5%; Score 7; DB 1; Length 33;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 446 EEDNDMD 452  
Db 232 EEDNDMD 238

Search completed: May 1, 2003, 20:53:44  
Job time : 44 secs

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:50:11 / Search time 89 Seconds

(without alignments)  
1053.387 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 455  
Sequence: 1 MSSLSGAFVQIKFDDLOFFE.....GDDDDDDDEEDNDMNSE 455

Scoring table: OLIGO

Gapop 60.0, Capext 60.0

Searched: 671580 seqs, 206047115 residues

Database size: 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database:

SPTREMBL\_21:\*

- 1: sp\_Archaea:\*
- 2: sp\_Bacteria:\*
- 3: sp\_Fungi:\*
- 4: sp\_Human:\*
- 5: sp\_Invertebrate:\*
- 6: sp\_Mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_Organelle:\*
- 9: sp\_Phage:\*
- 10: sp\_Plant:\*
- 11: sp\_Rodent:\*
- 12: sp\_Virus:\*
- 13: sp\_Vertebrate:\*
- 14: sp\_Unclassified:\*
- 15: sp\_Virus:\*
- 16: sp\_Bacteriophage:\*
- 17: sp\_Archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	455	100.0	455	4	Q9HCC4
2	331	72.7	800	4	Q9HDD2
3	331	72.7	800	4	Q9HCC5
4	331	72.7	800	4	Q9NY12
5	331	72.7	800	4	Q9NY12
6	166	36.5	454	11	Q9NY12
7	166	36.5	802	11	Q9ESL3
8	166	36.5	802	11	Q9ESL4
9	11	2.4	855	5	Q01700
10	11	2.4	977	5	Q9VW24
11	11	2.4	141	10	Q9SMA2
12	9	2.0	205	5	Q9SMA2
13	9	2.0	223	4	Q9SMA2
14	9	2.0	223	4	Q9SMA2
15	9	2.0	408	10	Q9SMA2
16	9	2.0	453	10	Q9SMA2

17	9	2.0	460	10	Q8SSG9	Q8SSG9 oryza sativ
18	9	2.0	502	11	Q91VY1	Q91VY1 mus musculu
19	9	2.0	502	11	Q91VY0	Q91VY0 mus musculu
20	9	2.0	504	11	Q63559	Q63559 rattus norv
21	9	2.0	607	10	Q9FG45	Q9FG45 arabidopsi
22	9	2.0	611	10	Q9LKB4	Q9LKB4 arabidopsi
23	9	2.0	616	5	Q9GRH2	Q9GRH2 sycon rapha
24	9	2.0	652	16	Q9A0J5	Q9A0J5 streptococ
25	9	2.0	764	3	Q12500	Q12500 saccharomyc
26	9	2.0	789	2	Q9XBP8	Q9XBP8 myxococcu
27	9	2.0	856	10	Q9ZSM8	Q9ZSM8 arabidopsi
28	9	2.0	898	10	Q04246	Q04246 arabidopsi
29	9	2.0	1059	3	Q9P8G2	Q9P8G2 bombyx mori
30	9	2.0	1472	5	Q9U5A8	Q9U5A8 anopheles g
31	9	2.0	1978	5	Q8TSH2	Q8TSH2 plasmodium
32	9	2.0	2391	5	Q27732	Q27732 plasmodium
33	8	1.8	101	4	Q15452	Q15452 homo sapien
34	8	1.8	106	12	Q91872	Q91872 spodoptera
35	8	1.8	157	12	Q55500	Q55500 human cytom
36	8	1.8	161	12	Q91902	Q91902 human cytom
37	8	1.8	161	12	Q91902	Q91902 human cytom
38	8	1.8	162	16	Q9HMO2	Q9HMO2 culicx nigri
39	8	1.8	180	3	Q9P840	Q9P840 pseudomonas
40	8	1.8	183	4	Q00193	Q00193 candida alb
41	8	1.8	184	10	Q42341	Q42341 arabidopsi
42	8	1.8	194	12	Q9QAM0	Q9QAM0 white spot
43	8	1.8	194	12	Q91LE7	Q91LE7 white spot
44	8	1.8	194	12	Q8QTE0	Q8QTE0 white spot
45	8	1.8	198	10	Q9C7Y9	Q9C7Y9 arabidopsi
46	8	1.8	204	10	Q93756	Q93756 tagus sylvia
47	8	1.8	206	10	Q9SM77	Q9SM77 hordeum vul
48	8	1.8	209	10	Q945F0	Q945F0 nicotiana t
49	8	1.8	211	17	Q28960	Q28960 archaeoglob
50	8	1.8	213	11	Q91VH4	Q91VH4 mus musculu
51	8	1.8	218	10	Q8R2R9	Q8R2R9 oryza sativ
52	8	1.8	229	16	Q8X4P6	Q8X4P6 escherichia
53	8	1.8	237	2	Q50439	Q50439 mycobacteri
54	8	1.8	246	10	Q9A590	Q9A590 oryza sativ
55	8	1.8	250	10	Q9AMH5	Q9AMH5 arabidopsi
56	8	1.8	258	5	Q9GR09	Q9GR09 leishmania
57	8	1.8	259	11	Q9CXN1	Q9CXN1 mus musculu
58	8	1.8	260	10	Q9S1S0	Q9S1S0 arabidopsi
59	8	1.8	277	10	Q948F2	Q948F2 oryza sativ
60	8	1.8	279	11	Q64125	Q64125 mus sp. ear
61	8	1.8	292	5	Q02445	Q02445 heliobact
62	8	1.8	293	4	Q9NMW4	Q9NMW4 homo sapien
63	8	1.8	297	5	Q17319	Q17319 caenorhabdi
64	8	1.8	298	10	Q64936	Q64936 arabidopsi
65	8	1.8	302	12	Q8VB23	Q8VB23 white spot
66	8	1.8	307	5	Q9NKR3	Q9NKR3 dirosophila
67	8	1.8	317	11	P97827	P97827 rattus norv
68	8	1.8	318	13	Q9PXX7	Q9PXX7 lecheron
69	8	1.8	318	17	Q27264	Q27264 methanobact
70	8	1.8	321	12	Q913B5	Q913B5 white spot
71	8	1.8	321	12	Q9VAL2	Q9VAL2 white spot
72	8	1.8	321	12	Q91L82	Q91L82 white spot
73	8	1.8	325	10	Q8S672	Q8S672 oryza sativ
74	8	1.8	326	10	Q8S672	Q8S672 oryza sativ
75	8	1.8	335	5	Q8TJF9	Q8TJF9 dictyosteli
76	8	1.8	335	4	Q9COP9	Q9COP9 homo sapien
77	8	1.8	334	4	Q9BXC3	Q9BXC3 homo sapien
78	8	1.8	359	10	Q8RXA9	Q8RXA9 zea mays (m
79	8	1.8	359	5	Q26776	Q26776 tyrannosoma
80	8	1.8	368	10	Q8VZN7	Q8VZN7 arabidopsi
81	8	1.8	368	10	Q8SGW5	Q8SGW5 arabidopsi
82	8	1.8	374	5	Q76381	Q76381 caenorhabdi
83	8	1.8	379	10	Q49391	Q49391 arabidopsi
84	8	1.8	380	10	Q91V94	Q91V94 arabidopsi
85	8	1.8	395	11	Q9ULD4	Q9ULD4 rattus norv
86	8	1.8	396	11	Q923V6	Q923V6 rattus norv
87	8	1.8	414	5	Q8T030	Q8T030 dirosophila
88	8	1.8	415	10	Q9MAV2	Q9MAV2 arabidopsi
89	8	1.8	416	4	Q9P289	Q9P289 homo sapien

90 8 1.8 416 11 Q99J72 Q99J72 mus musculu  
 91 8 1.8 422 10 Q40543 Q40543 nicotiana t  
 92 8 1.8 426 10 Q95ZM8 Q95ZM8 arabidopsis  
 93 8 1.8 434 5 Q23994 Q23994 drosophila  
 94 8 1.8 434 13 Q91B76 Q91B76 pagrus majo  
 95 8 1.8 436 10 Q91FV3 Q91FV3 arabidopsis  
 96 8 1.8 438 10 Q9M3H8 Q9M3H8 cicer ariet  
 97 8 1.8 439 5 O16303 O16303 caenorhabd  
 98 8 1.8 443 13 Q91863 Q91863 fugu rubrid  
 99 8 1.8 445 11 Q60877 Q60877 mus musculu  
 100 8 1.8 447 13 Q8QCV6 Q8QCV6 xenopus lae

## ALIGNMENTS

RESULT 1  
 ID Q9HCC4 PRELIMINARY; PRT; 455 AA.

01-MAR-2001 (TrEMBLrel. 16, Created)  
 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 MURK-beta (Similar to sterile-alpha motif and leucine zipper  
 containing kinase AKZ) (Mixed lineage kinase) (Mixed lineage  
 kinase-related kinase MRK-beta).  
 MLTK.  
 Homo sapiens (Human).  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NCBI\_Taxid=9606;  
 [1]  
 SEQUENCE FROM N.A.  
 Gotoh T., Adachi M., Nishida E.;  
 Identification and characterization of a novel MAP kinase Kinase  
 Kinase, MLTK."  
 Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 [2]  
 SEQUENCE FROM N.A.  
 TISSUE-COLON;  
 Strauberg R.;  
 Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 [3]  
 SEQUENCE FROM N.A.  
 Acton S.;  
 "MLK-mixed lineage kinase.";  
 Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 [4]  
 SEQUENCE FROM N.A.  
 MEDLINE=21950776; PubMed=11836244;  
 Gross E.A., Callow M.G., Waldbaum L., Thomas S., Ruggieri R.;  
 "MRK, a Mixed lineage kinase-related Molecule that Plays a Role in  
 Gamma-Radiation-induced Cell Cycle Arrest.";  
 J. Biol. Chem. 277:13873-13882(2002).  
 -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 EMBL; AB049734; BAB16445.1; -  
 EMBL; BC001401; AA011401.1; -  
 EMBL; AF325454; AKI1615.1; -  
 EMBL; AF480462; AAL85892.1; -  
 HSSP; P12931; IFMK.  
 InterPro: IPR000719; Euk\_Pkinase.  
 InterPro: IPR002290; Ser\_thr\_kinase.  
 InterPro: IPR004040; STY\_kinase.  
 InterPro: IPR001245; Tyr\_kinase.  
 Pfam; PF00069; Pkinase; 1.  
 PRINTS; PR00109; TYRKINASE.  
 ProDom; PD000001; Euk\_Pkinase; 1.  
 SMART; SM00221; STYK; 1.  
 SMART; SM00220; S\_TK; 1.  
 SMART; SM00219; TYK; 1.  
 PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.

SQ SEQUENCE 455 AA; 51582 MW; E87DB84A4D58B752 CRC64;

Query Match 100.0%; Score 455; DB 4; Length 455;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSIGASVQIKPDLDLFFENCGGSGFGSYRAKMWISQDKEVAVKLLKIEKAELISVL 60  
 1 MSSIGASVQIKPDLDLFFENCGGSGFGSYRAKMWISQDKEVAVKLLKIEKAELISVL 60  
 DB 1 MSSIGASVQIKPDLDLFFENCGGSGFGSYRAKMWISQDKEVAVKLLKIEKAELISVL 60  
 QY 61 SRRNIQFYGVILPEPNVGIETEVASLGSIDYDINSRSEMDMDHMTATVAKGMV 120  
 61 SRRNIQFYGVILPEPNVGIETEVASLGSIDYDINSRSEMDMDHMTATVAKGMV 120  
 DB 61 SRRNIQFYGVILPEPNVGIETEVASLGSIDYDINSRSEMDMDHMTATVAKGMV 120  
 QY 121 LHMEAPVYVTHDLSKRVVIAADGVLTICDPGASRFHNTHTMSLVGTFPMNAPEVIQS 180  
 121 LHMEAPVYVTHDLSKRVVIAADGVLTICDPGASRFHNTHTMSLVGTFPMNAPEVIQS 180  
 DB 121 LHMEAPVYVTHDLSKRVVIAADGVLTICDPGASRFHNTHTMSLVGTFPMNAPEVIQS 180  
 QY 181 LPSVETCDTYSYGVVLAAMLTREVPFKGLBGLQVAMLVENRERLTPSSCPSPFALLH 240  
 181 LPSVETCDTYSYGVVLAAMLTREVPFKGLBGLQVAMLVENRERLTPSSCPSPFALLH 240  
 DB 181 LPSVETCDTYSYGVVLAAMLTREVPFKGLBGLQVAMLVENRERLTPSSCPSPFALLH 240  
 QY 241 QCEWADAKRRPSFKQIISLESMSNDTSLPDKNSFLHNKAEMRCIEATLERLKLIERD 300  
 241 QCEWADAKRRPSFKQIISLESMSNDTSLPDKNSFLHNKAEMRCIEATLERLKLIERD 300  
 DB 241 QCEWADAKRRPSFKQIISLESMSNDTSLPDKNSFLHNKAEMRCIEATLERLKLIERD 300  
 QY 301 LSPFEQELKERERRRLKMEQKLTQESNTPLPLAARMSSESYESTKTESNSAEMSCOI 360  
 301 LSPFEQELKERERRRLKMEQKLTQESNTPLPLAARMSSESYESTKTESNSAEMSCOI 360  
 DB 301 LSPFEQELKERERRRLKMEQKLTQESNTPLPLAARMSSESYESTKTESNSAEMSCOI 360  
 QY 361 TATSNBEGHGMPSIQAMLMGFGDIFSNKAGAVHSGMOINMOAKNSSKTTSKRGGK 420  
 361 TATSNBEGHGMPSIQAMLMGFGDIFSNKAGAVHSGMOINMOAKNSSKTTSKRGGK 420  
 DB 361 TATSNBEGHGMPSIQAMLMGFGDIFSNKAGAVHSGMOINMOAKNSSKTTSKRGGK 420  
 QY 421 KVMALGFSDPDLSEGGDDDDDDGEEEDNDMDNSE 455  
 421 KVMALGFSDPDLSEGGDDDDDDGEEEDNDMDNSE 455  
 DB 421 KVMALGFSDPDLSEGGDDDDDDGEEEDNDMDNSE 455

## RESULT 2

ID Q9HDD2 PRELIMINARY; PRT; 800 AA.

01-MAR-2001 (TrEMBLrel. 16, Created)  
 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 Placible mixed-lineage kinase protein.  
 MLTK.  
 Homo sapiens (Human).  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NCBI\_Taxid=9606;  
 [1]  
 SEQUENCE FROM N.A.  
 TISSUE-LYMPHOID ORGAN;  
 Abe Y., Ueda N.;  
 "Placible Mixed-lineage kinase derived from LAK cell.";  
 Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 EMBL; AB030034; BAB12040.1; -  
 HSSP; P12931; IFMK.  
 InterPro: IPR000719; Euk\_Pkinase.  
 InterPro: IPR001660; SAM.  
 InterPro: IPR002290; Ser\_thr\_kinase.  
 InterPro: IPR004040; STY\_kinase.  
 InterPro: IPR001245; Tyr\_kinase.  
 Pfam; PF00069; Pkinase; 1.  
 PRINTS; PR00109; TYRKINASE.  
 ProDom; PD000001; Euk\_Pkinase; 1.  
 SMART; SM00221; STYK; 1.  
 SMART; SM00220; S\_TK; 1.  
 SMART; SM00219; TYK; 1.

DR SMART: SM00219; TYKc; 1.  
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 KM ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 800 AA; 91155 MW; B2814509EC54B07A CRC64;

Query Match 72.7%; Score 331; DB 4; Length 800;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFPDLOPFENCGGSGFSYVRAMISQDKEVAVKLLIKEAEIISVL 60  
 Db 1 MSSLGASFVQIKFPDLOPFENCGGSGFSYVRAMISQDKEVAVKLLIKEAEIISVL 60  
 QY 61 SHRNIIQFYGVILPPNYGIVTEYASIGSLYDYNRSRSEMDMHIIMTATVAKGMHY 120  
 Db 61 SHRNIIQFYGVILPPNYGIVTEYASIGSLYDYNRSRSEMDMHIIMTATVAKGMHY 120  
 QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTTHMSLVGTFPMAPEVIQS 180  
 Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTTHMSLVGTFPMAPEVIQS 180  
 QY 181 LPVSETCDTYSYGVVLEMLTREVPFKGLEGLQVAMLVYKNERLTIPSSCPSPFALLH 240  
 Db 181 LPVSETCDTYSYGVVLEMLTREVPFKGLEGLQVAMLVYKNERLTIPSSCPSPFALLH 240  
 QY 241 QCWEADAKKRPSPFKQIISIESMSNDTSLPDKCNSEFLHNKAERCEIATLERLTKLERD 300  
 Db 241 QCWEADAKKRPSPFKQIISIESMSNDTSLPDKCNSEFLHNKAERCEIATLERLTKLERD 300  
 QY 301 LSPFEQELKERERRLKMEQKLTQGSNTPL 331  
 Db 301 LSPFEQELKERERRLKMEQKLTQGSNTPL 331

RESULT 3  
 Q9HCC5 PRELIMINARY; PRT; 800 AA.

AC Q9HCC5; 01-OCT-2000 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE MLTK-alpha.  
 GN MLTK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NCBI\_TaxID=9606;  
 (1)  
 SEQUENCE FROM N.A.  
 RX MEDLINE=2126427; PubMed=11042189;  
 RA Gotch I., Adachi M., Nishida E.,  
 RT "Identification and Characterization of a Novel MAP Kinase Kinase  
 RT Kinase, MLTK.";  
 RL J. Biol. Chem. 276:4276-4286(2001).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AB049733; BAB16444.1; -.  
 DR HSSP; P12931; IFMK.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001660; SAM.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00536; SAM; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00454; SAM; 1.  
 DR SMART; SM00221; STYK; 1.  
 DR SMART; SM00219; STYK; 1.  
 DR SMART; SM00219; TYKc; 1.  
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.

KM ATP-binding; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 800 AA; 91188 MW; 2C8593824AB3FADD CRC64;

Query Match 72.7%; Score 331; DB 4; Length 800;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFPDLOPFENCGGSGFSYVRAMISQDKEVAVKLLIKEAEIISVL 60  
 Db 1 MSSLGASFVQIKFPDLOPFENCGGSGFSYVRAMISQDKEVAVKLLIKEAEIISVL 60  
 QY 61 SHRNIIQFYGVILPPNYGIVTEYASIGSLYDYNRSRSEMDMHIIMTATVAKGMHY 120  
 Db 61 SHRNIIQFYGVILPPNYGIVTEYASIGSLYDYNRSRSEMDMHIIMTATVAKGMHY 120  
 QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTTHMSLVGTFPMAPEVIQS 180  
 Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFHNTTHMSLVGTFPMAPEVIQS 180  
 QY 181 LPVSETCDTYSYGVVLEMLTREVPFKGLEGLQVAMLVYKNERLTIPSSCPSPFALLH 240  
 Db 181 LPVSETCDTYSYGVVLEMLTREVPFKGLEGLQVAMLVYKNERLTIPSSCPSPFALLH 240  
 QY 241 QCWEADAKKRPSPFKQIISIESMSNDTSLPDKCNSEFLHNKAERCEIATLERLTKLERD 300  
 Db 241 QCWEADAKKRPSPFKQIISIESMSNDTSLPDKCNSEFLHNKAERCEIATLERLTKLERD 300  
 QY 301 LSPFEQELKERERRLKMEQKLTQGSNTPL 331  
 Db 301 LSPFEQELKERERRLKMEQKLTQGSNTPL 331

RESULT 4  
 Q9NYL2 PRELIMINARY; PRT; 800 AA.

AC Q9NYL2; 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Mixed lineage kinase.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NCBI\_TaxID=9606;  
 (1)  
 SEQUENCE FROM N.A.  
 RX MEDLINE=20384179; PubMed=10924358;  
 RA Liu T.C., Huang C.J., Chu Y.C., Wei C.C., Chou C.C., Chou M.Y.,  
 RA Chou C.K., Yang J.J.;  
 RT "Cloning and expression of ZAK, a mixed lineage kinase-like protein  
 RT containing a leucine-zipper and a sterile-alpha motif.";  
 RL Biochem. Biophys. Res. Commun. 274:811-816(2000).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AF238255; AA63490.1; -.  
 DR HSSP; P12931; IFMK.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR001660; SAM.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR004040; STY\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00536; SAM; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00454; SAM; 1.  
 DR SMART; SM00221; STYK; 1.  
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 KM ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 800 AA; 91264 MW; DA82D7ABB2082F43 CRC64;

Query Match 72.7%; Score 331; DB 4; Length 800;  
 Best Local Similarity 100.0%; Pred. No. 0;

Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSIGASFOVQIKFDDLOPFENCGGSGFSYVRANKISODKEVAVKLLIKKEAEILLSVL 60  
 Db 1 MSSIGASFOVQIKFDDLOPFENCGGSGFSYVRANKISODKEVAVKLLIKKEAEILLSVL 60

QY 61 SHRNIIQFYGVILEPPNYGIYEVASLGSLYDYINSNSEEMDMHIMTATVAKGMHY 120  
 Db 61 SHRNIIQFYGVILEPPNYGIYEVASLGSLYDYINSNSEEMDMHIMTATVAKGMHY 120

QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFNNHTHMSLVGTFPMAPEVIQS 180  
 Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFNNHTHMSLVGTFPMAPEVIQS 180

QY 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEGIQVAMLYVEKNERLTISSCPSPFAELH 240  
 Db 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEGIQVAMLYVEKNERLTISSCPSPFAELH 240

QY 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEWRCIEATLERLKKLERD 300  
 Db 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEWRCIEATLERLKKLERD 300

QY 301 LSPKEQELKERERRRLKMWEOKLTQESNTPL 331  
 Db 301 LSPKEQELKERERRRLKMWEOKLTQESNTPL 331

RESULT 5

Q9NVE9 PRELIMINARY; PRT; 800 AA.

ID Q9NVE9  
 AC Q9NVE9;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Sterile-alpha motif and leucine zipper containing kinase AZK (Mixed lineage kinase-related kinase MRK-alpha).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA McNeé J.J., Frima N., Diamond T.E., Dower S.K., Guesdon F.;  
 RT "Cloning and characterization of AZK, a mixed lineage kinase containing a sterile-alpha motif."  
 RT Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.  
 RL [2]  
 CC EMBL: AR251441; AAF65822.1; -;  
 DR EMBL: AF480461; AAL85891.1; -;  
 DR HSSP: P12931; IFMK.  
 DR InterPro: IPR000719; Btk\_pkinase.  
 DR InterPro: IPR001660; SAM.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR004040; STY\_pkinase.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00069; pkinase; 1.  
 DR Pfam: PF00536; SAM; 1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR ProDom: PD000001; Btk\_pkinase; 1.  
 DR SMART: SM00454; SAM; 1.  
 DR SMART: SM00221; STYK; 1.  
 DR PROSITE: PSS0011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 KW ATP-binding; 800 AA; 91181 MW; B289D836E5C5E295 CRC64;  
 SQ SEQUENCE

Query Match 72.7%; Score 331; DB 4; Length 800;  
 Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0;  
 Matches 331; Conservative 0; Indels 0; Gaps 0;

QY 1 MSSIGASFOVQIKFDDLOPFENCGGSGFSYVRANKISODKEVAVKLLIKKEAEILLSVL 60  
 Db 1 MSSIGASFOVQIKFDDLOPFENCGGSGFSYVRANKISODKEVAVKLLIKKEAEILLSVL 60

QY 61 SHRNIIQFYGVILEPPNYGIYEVASLGSLYDYINSNSEEMDMHIMTATVAKGMHY 120  
 Db 61 SHRNIIQFYGVILEPPNYGIYEVASLGSLYDYINSNSEEMDMHIMTATVAKGMHY 120

QY 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFNNHTHMSLVGTFPMAPEVIQS 180  
 Db 121 LHMEAPVKVIHRDLKSRNVVIAADGVLCIDFGASRFNNHTHMSLVGTFPMAPEVIQS 180

QY 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEGIQVAMLYVEKNERLTISSCPSPFAELH 240  
 Db 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEGIQVAMLYVEKNERLTISSCPSPFAELH 240

QY 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEWRCIEATLERLKKLERD 300  
 Db 241 QCWEADAKRPSFKQIISILESMSNDTSLPDKNSFLHNKAEWRCIEATLERLKKLERD 300

QY 301 LSPKEQELKERERRRLKMWEOKLTQESNTPL 331  
 Db 301 LSPKEQELKERERRRLKMWEOKLTQESNTPL 331

RESULT 6

Q9ESL3 PRELIMINARY; PRT; 454 AA.

ID Q9ESL3  
 AC Q9ESL3;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE MLTK-beta.  
 GN ZAK OR MLTK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA McNeé J.J., Frima N., Diamond T.E., Dower S.K., Guesdon F.;  
 RT "Cloning and characterization of a Novel MAP Kinase Kinase Kinase, MLTK."  
 RT Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.  
 RL J. Biol. Chem. 276:4276-4286(2001).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL: AB049732; BAB16443.1; -;  
 DR HSSP: P12931; IFMK.  
 DR MGD: MGI:1931274; Zak.  
 DR InterPro: IPR000719; Btk\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR004040; STY\_pkinase.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00069; pkinase; 1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR ProDom: PD000001; Btk\_pkinase; 1.  
 DR SMART: SM00221; STYK; 1.  
 DR SMART: SM00220; S\_TYK; 1.  
 DR SMART: SM00219; TYRK; 1.  
 DR PROSITE: PSS0011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR ATP-binding; Serine/threonine-protein kinase; Transferase.  
 KW ATP-binding; 454 AA; 51366 MW; 35C2PFC0D729D9395 CRC64;  
 SQ SEQUENCE

Query Match 36.5%; Score 166; DB 11; Length 454;  
 Best Local Similarity 99.6%; Pred. No. 3; e-168;  
 Matches 266; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSIGASFOVQIKFDDLOPFENCGGSGFSYVRANKISODKEVAVKLLIKKEAEILLSVL 60

```

Db 1 MSSLGASFWQIKFDLQFFENCGGSGFSGYVRAKWI SODEKVAVKLLKIKKAEKAILSVL 60
Qy 61 SHRNIIQFYGVILEPPNYGIVTEYASLGSLYDINSRSEMDHIMTATDVAKGMHY 120
Db 61 SHRNIIQFYGVILEPPNYGIVTEYASLGSLYDINSRSEMDHIMTATDVAKGMHY 120
Qy 121 LHMEAPVAVIHRDLKSRNVVIAADGVLCIDFGASRFNNHTTHMSLVGTFPMMAPEVIQS 180
Db 121 LHMEAPVAVIHRDLKSRNVVIAADGVLCIDFGASRFNNHTTHMSLVGTFPMMAPEVIQS 180
Qy 181 LPVSETCDTYSYGVLMEMLTREVPFGKLGLOVAVLVKNERLTPSSCPRSPFAELH 240
Db 181 LPVSETCDTYSYGVLMEMLTREVPFGKLGLOVAVLVKNERLTPSSCPRSPFAELH 240
Qy 241 QCWEADAKKRPSPFKQIISILESMSNDT 267
Db 241 QCWEADAKKRPSPFKQIISILESMSNDT 267

RESULT 7
QESL4 PRELIMINARY; PRT; 802 AA.
ID Q9ESL4;
AC Q9ESL4;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE MLTK alpha.
GN ZAK OR MLTK.
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21264927; PubMed=11042189;
RA Gochi I., Adachi M., Nishida E.;
RT "Identification and Characterization of a Novel MAP Kinase Kinase
RT Kinase, MLTK."
RT J. Biol. Chem. 276:4276-4286(2001);
CC -1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB049731; BAB16442.1; -.
DR HSSP; P12931; IFMK.
DR MCD; MGI:1931274; ZAK.
DR InterPro; IPR000194; ATPase_a/bcentre.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; Str_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Euk_pkinase; 1.
DR SMART; SM00454; SAM; 1.
DR SMART; SM00221; STYKC; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TYKC; 1.
DR PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
DR PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 802 AA; 91719 MW; D431DF8312A43CC CRC64;

Query Match 36.5%; Score 166; DB 11; Length 802;
Best Local Similarity 99.6%; Pred. NO. 5.6e-168;
Matches 266; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSSLGASFWQIKFDLQFFENCGGSGFSGYVRAKWI SODEKVAVKLLKIKKAEKAILSVL 60
Db 1 MSSLGASFWQIKFDLQFFENCGGSGFSGYVRAKWI SODEKVAVKLLKIKKAEKAILSVL 60
Qy 61 SHRNIIQFYGVILEPPNYGIVTEYASLGSLYDINSRSEMDHIMTATDVAKGMHY 120

```

```

Db 61 SHRNIIQFYGVILEPPNYGIVTEYASLGSLYDINSRSEMDHIMTATDVAKGMHY 120
Qy 121 LHMEAPVAVIHRDLKSRNVVIAADGVLCIDFGASRFNNHTTHMSLVGTFPMMAPEVIQS 180
Db 121 LHMEAPVAVIHRDLKSRNVVIAADGVLCIDFGASRFNNHTTHMSLVGTFPMMAPEVIQS 180
Qy 181 LPVSETCDTYSYGVLMEMLTREVPFGKLGLOVAVLVKNERLTPSSCPRSPFAELH 240
Db 181 LPVSETCDTYSYGVLMEMLTREVPFGKLGLOVAVLVKNERLTPSSCPRSPFAELH 240
Qy 241 QCWEADAKKRPSPFKQIISILESMSNDT 267
Db 241 QCWEADAKKRPSPFKQIISILESMSNDT 267

RESULT 8
Q90ZY8 PRELIMINARY; PRT; 371 AA.
ID Q90ZY8;
AC Q90ZY8;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Protein kinase Npk.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxId=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Chou C.-M., Lee I.-L., Leu J.-H., Huang C.-J.;
RT "A novel protein kinase, znpk, from the zebrafish."
RT Submitted (MAY-2000) to the EMBL/Genbank/DBD databases.
DR EMBL; AF265343; AAK52416.1; -.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRODOM; PD000001; Euk_pkinase; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 371 AA; 42456 MW; 9B918B8AB820D296 CRC64;

Query Match 10.8%; Score 49; DB 13; Length 371;
Best Local Similarity 100.0%; Pred. NO. 2.6e-43;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 160 HTTMSLVGTFPMMAPEVIQS LPVSETCDTYSYGVLMEMLTREVPFGK 208
Db 191 HTTMSLVGTFPMMAPEVIQS LPVSETCDTYSYGVLMEMLTREVPFGK 239

RESULT 9
001700 PRELIMINARY; PRT; 855 AA.
ID 001700;
AC 001700;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Hypothetical 95.1 kDa protein F33E2.2.
GN F33E2.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Leonard N.;
RT Submitted (JAN-1997) to the EMBL/Genbank/DBD databases.
CC -1 SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AL022593; CAA18635.2; -.
DR EMBL; Z84574; CAA18635.2; JOINED.

```

DR EMBL: Z84574: CAB06544.2; -.  
 DR EMBL: AL022593: CAB06544.2; JOINED.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR004040; STY\_pkinase.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00069; pkinase.1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR ProDom: PD000001; Euk\_pkinase.1.  
 DR SMART: SM00221; STYKC.1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR ATP-binding; Hypothetical protein; Serine/threonine-protein kinase;  
 KW Transferase.  
 SQ SEQUENCE 855 AA; 95786 MW; P44D2538CB7D95A CRC64;

Query Match  
 Best Local Similarity 100.0%; Pred. No. 0.021;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

193 GVLMEMTLRE 203  
 |||||  
 235 GVLMEMTLRE 245

RESULT 10  
 ID Q9VM24 PRELIMINARY; PRT; 977 AA.  
 AC Q9VM24.  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE CG8789 protein (Drosophila).  
 OS CG8789.  
 NC Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.P.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Abmayyan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,  
 RA Borovica D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Paulis B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Folsler K., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
 RA Glisdek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck M.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jajani M., Kalish F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Jimali B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshireli A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,  
 RA Palazolo M., Pittman K.G., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RA Stapleton M., Brokstein P., Hong L., Abmayyan A., Carlson J.,  
 RA Change M., Chavez C., Dorsett V., Dresnak D., Farfan D., Fiske E.,  
 RA George R., Gonzalez M., Guarin H., Kromoller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.U., Nurno J., Pacle J., Paragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Celisner S.;  
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL: AF094787; AF094787.1; -.  
 DR EMBL: AF094787; AF094787.1; -.  
 DR HSSP: P08631; IAD5.  
 DR FlyBase: FBgn0036896; CG8789.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR004040; STY\_pkinase.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00069; pkinase.1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR ProDom: PD000001; Euk\_pkinase.1.  
 DR SMART: SM00221; STYKC.1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 977 AA; 109567 MW; 5CD23593C014BD4E CRC64;

Query Match  
 Best Local Similarity 100.0%; Pred. No. 0.024;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 191 SYGVLMEMTL 201  
 |||||  
 DB 331 SYGVLMEMTL 341

RESULT 11  
 ID Q9SMA2 PRELIMINARY; PRT; 141 AA.  
 AC Q9SMA2.  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE Zm00007.1.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriophytidae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. INDICA;  
 RA Hong G., Zhao W.;  
 RT "Oryza sativa genomic DNA, chromosome 4, clone: b6015";  
 RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL117264; CAB55391.1; -.  
 SQ SEQUENCE 141 AA; 15696 MW; 1C608AE0459DD77D CRC64;

Query Match  
 Best Local Similarity 100.0%; Pred. No. 0.56;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 439 DDDDDGEEE 447  
 |||||



DB 26 DDDDDGEE 34

## RESULT 12

Q8ST12

PRELIMINARY; PRT; 205 AA.

DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DE Hypothetical protein. 6/101.  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
 NCBI\_TaxId=44689;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AX4;  
 RA Gloeckner G., Eichinger L., Szafrenski K., Pachbat J., Dear P.,  
 Lehmann R., Baugart C., Parra G., April J.F., Guigo R., Kumpf K.,  
 Tunggal B., Cox E., Quail M.A., Platzer W., Rosenthal A., Noegel A.A.;  
 "Sequence and Analysis of Chromosome 2 of Dictyostelium.";  
 Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; ACl15578; AAL92211.1;  
 KW Hypothetical protein.  
 SQ SEQUENCE 205 AA; 22960 MW; 7E3D24AD2290C4D0 CRC64;

Query Match  
 Best Local Similarity 2.0%; Score 9; DB 5; Length 205;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 EGDGDDDDDD 443  
 DB 56 EGDGDDDDDD 64

## RESULT 13

Q15318

PRELIMINARY; PRT; 233 AA.

DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DE RNC polymerase III subunit.  
 GN RNC32.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=97315201; PubMed=9171375;  
 RA Wang Z., Roeder R.G.;  
 RT "Three human RNA polymerase III-specific subunits form a subcomplex  
 with a selective function in specific transcription initiation.";  
 RL Genes Dev. 11:1335-1326(1997).  
 DR EMBL; U93688; AAB63676.1;  
 SQ SEQUENCE 233 AA; 27299 MW; 16EAB6AA50002PF CRC64;

Query Match  
 Best Local Similarity 2.0%; Score 9; DB 4; Length 233;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 EGDGDDDDDD 443  
 DB 173 EGDGDDDDDD 181

## RESULT 14

Q98182

PRELIMINARY; PRT; 297 AA.

DT 01-FEB-1997 (TREMBlrel. 02, Created)  
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE MC011L.  
 GN MC011L.  
 OS Molluscum contagiosum virus subtype 1 (MCV1).  
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Molluscipoxvirus.  
 NCBI\_TaxId=10280;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=96325459; PubMed=8670425;  
 RA Senkevich T.G., Bugert J.J., Sialer J.R., Koonin E.V., Darai G.,  
 Moss B.;  
 RT "Genome sequence of a human tumorigenic poxvirus: Prediction of  
 specific host response-evasion genes.";  
 RL Science 273:813-816(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC Senkevich T.G., Bugert J.J., Sialer J.R., Koonin E.V., Darai G.,  
 Moss B.;  
 RA Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U60315; AAC55139.1;  
 DR InterPro; IPR001230; Ptenyl site.  
 DR PROSITE; PS00294; PRENYLATION; UNKNOWN 1.  
 SQ SEQUENCE 297 AA; 31500 MW; 86199E36CD8A781 CRC64;

Query Match  
 Best Local Similarity 2.0%; Score 9; DB 12; Length 297;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDG 444  
 DB 206 GDDDDDDG 214

## RESULT 15

Q39585

PRELIMINARY; PRT; 408 AA.

DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DE Protein kinase.  
 GN Chlamydomonas reinhardtii.  
 OS Chlamydomonas reinhardtii.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 NCBI\_TaxId=3055;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=21GR(+);  
 RA MEDLINE=9613873; PubMed=8552645;  
 RA Kurvari V., Zhang Y., Luo Y., Snell W.J.;  
 RT "Molecular cloning of a protein kinase whose phosphorylation is  
 regulated by genetic adhesion during Chlamydomonas fertilization.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:39-43(1996).  
 DR EMBL; U36196; AAA96956.1;  
 DR InterPro; IPR000644; CBS\_domain.  
 DR Pfam; PF00571; CBS; 3.  
 DR SMART; SM00116; CBS; 3.  
 KW Kinase.  
 SQ SEQUENCE 408 AA; 44946 MW; DBDE6BB30C49CF2 CRC64;

Query Match  
 Best Local Similarity 2.0%; Score 9; DB 10; Length 408;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 EGDGDDDDDD 443  
 DB 393 EGDGDDDDDD 401

## RESULT 16

Q85SD1

PRELIMINARY; PRT; 453 AA.

AC 0855D1;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Hypothetical 49.5 kDa protein.  
 GN O01341F06.5.  
 OS Oryza sativa (Japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eubacteriales; Oryzaceae; Oryza.  
 RX NCBI\_Taxid=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV, NIPPONBARE;  
 RA McCombie W.R., Spiegel L., de la Bastide M., Preston R., Ferraro K.,  
 RA Kuit K., Nascimento L., Zutterer T., Ballia V., Bell M., Baker J.,  
 RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,  
 RA Dike S., O'Shaughnessy A., Palmer L., Dedhia N.,  
 RA "Genomic sequence for Oryza sativa, Nipponbare strain, clone  
 OJ1341F06, from chromosome 10, complete sequence."  
 Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 EMBL; AC116926; AAM0885.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 453 AA; 49540 MW; 3661B92B2399644 CRC64;  
 Query Match 2.0%; Score 9; DB 10; Length 453;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 436 GDDDDDDG 444  
 DB 277 GDDDDDDG 285

RESULT 17  
 ID 0855G9 PRELIMINARY; PRT; 460 AA.  
 AC 0855G9;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Hypothetical 50.4 kDa protein.  
 GN OSUNBA0091J06.10.  
 OS Oryza sativa (Japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eubacteriales; Oryzaceae; Oryza.  
 RX NCBI\_Taxid=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV, NIPPONBARE;  
 RA McCombie W.R., de la Bastide M., Spiegel L., Preston R., Ferraro K.,  
 RA Kuit K., Nascimento L., Zutterer T., Ballia V., Bell M., Baker J.,  
 RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,  
 RA Dike S., O'Shaughnessy A., Palmer L., Dedhia N.,  
 RA "Genomic sequence for Oryza sativa, Nipponbare strain, clone  
 OSUNBA0091J06, from chromosome 10, complete sequence."  
 Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 EMBL; AC113338; AAM08654.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 460 AA; 50423 MW; 4624B7D90845941A CRC64;  
 Query Match 2.0%; Score 9; DB 10; Length 460;  
 Best Local Similarity 100.0%; Pred. No. 1.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 436 GDDDDDDG 444  
 DB 284 GDDDDDDG 292

RESULT 18  
 Q91YV1

ID Q91YV1 PRELIMINARY; PRT; 502 AA.  
 AC Q91YV1;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Similar to activin A receptor, type II-like 1.  
 GN ACVRL1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Strausberg R.;  
 RA Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 EMBL; BC014291; AAH14291.1; -  
 DR MGD; MGI:1338946; Acvrl1.  
 DR InterPro: IPR000472; Activin\_rec.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF01064; Activin\_rec1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 DR PROSITE; PS00101; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
 KW ATP-binding; Receptor; Transferase.  
 SQ SEQUENCE 502 AA; 56483 MW; 800E8254703AE875 CRC64;  
 Query Match 2.0%; Score 9; DB 11; Length 502;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 431 HRDLKSRNV 139  
 DB 327 HRDLKSRNV 335

RESULT 19  
 ID Q91YR0 PRELIMINARY; PRT; 502 AA.  
 AC Q91YR0;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Activin A receptor, type II-like 1.  
 GN ACVRL1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Strausberg R.;  
 RA Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 EMBL; BC015083; AAH15083.1; -  
 DR MGD; MGI:1338946; Acvrl1.  
 DR InterPro: IPR000472; Activin\_rec.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR Pfam; PF01064; Activin\_rec1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 DR PROSITE; PS00101; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; UNKNOWN\_1.  
 KW ATP-binding; Receptor; Transferase.  
 SQ SEQUENCE 502 AA; 56519 MW; 439510D3CCT740D65 CRC64;  
 Query Match 2.0%; Score 9; DB 11; Length 502;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 HRDLKSRNV 139  
 Db 327 HRDLKSRNV 335

## RESULT 20

ID 063559 PRELIMINARY; PRT; 504 AA.

AC 063559; 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Serine/threonine kinase receptor.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 KA MEDLINE=SPRAGUE-DAWLEY, TISSUE=LUNG;  
 KA MEDLINE=96198306; PubMed=8928814;  
 KA Panchenko M.P., Williams M.C., Brody J.S., Yu Q.;  
 RT "Type I receptor serine-threonine kinase preferentially expressed in  
 RT pulmonary blood vessels."  
 RL Am. J. Physiol. 270:L547-L558(1996).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL: I36088; AAC37705.1.  
 DR InterPro: IPR000472; Activin\_rec.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR003605; TGFbeta\_GS.  
 DR Pfam: PF01064; Activin\_recpt.1.  
 DR Pfam: PF00069; pkinase.1.  
 DR ProDom: PD000001; Euk\_pkinase.1.  
 DR SMART: SM00467; GS.1.  
 DR PROSITE: PS00107; PROTEIN KINASE ATP.1.  
 DR PROSITE: PS00101; PROTEIN KINASE DOM.1.  
 DR PROSITE: PS00108; PROTEIN KINASE ST.1.  
 DR ATP-binding: Serine/threonine-protein kinase; Transferase.  
 SO SEQUENCE 504 AA; 56703 MW; B40EA30775223C8F CRC64;

Query Match 2.0%; Score 9; DB 11; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 131 HRDLKSRNV 139  
 Db 327 HRDLKSRNV 337

## RESULT 21

ID 09FG45 PRELIMINARY; PRT; 607 AA.

AC 09FG45; 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Similarity to unknown protein.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 KA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,  
 RT Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI."  
 RT Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP000607; BAB10964.1; -.  
 SO SEQUENCE 607 AA; 70546 MW; 6A9CA9F08FF94634 CRC64;

Query Match 2.0%; Score 9; DB 10; Length 607;  
 Best Local Similarity 100.0%; Pred. No. 2.1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 436 GDDDDDDG 444  
 Db 586 GDDDDDDG 594

## RESULT 22

ID 09LKB4 PRELIMINARY; PRT; 611 AA.

AC 09LKB4; 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Genomic DNA, chromosome 3, TAC clone:K15M2.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 KA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;  
 RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 KA MEDLINE=20363099; PubMed=10907853;  
 KA Nakamura Y.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.  
 RT TAC and BAC clones."  
 RT DNA Res. 7:217-221(2000).  
 DR EMBL: AP000370; BAA97053.1; -.  
 SO SEQUENCE 611 AA; 71221 MW; ACE12328A9429778 CRC64;

Query Match 2.0%; Score 9; DB 10; Length 611;  
 Best Local Similarity 100.0%; Pred. No. 2.1;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 437 DDDDDDDG 445  
 Db 446 DDDDDDDG 454

## RESULT 23

ID 09GRH2 PRELIMINARY; PRT; 616 AA.

AC 09GRH2; 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Tyrosine Kinase.  
 OS Sycon raphanus.  
 OC Eukaryota; Metazoa; Porifera; Calcarea; Calcarea; Leucosolenida;  
 OC Sycetidae.  
 NCBI\_TaxID=56443;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 KA Skorokhod A., Gamulin V., Gundacker D., Kavan V., Mueller I.M.,  
 RT Mueller W.E.G.;  
 RT "Origin of insulin receptor tyrosine kinases in marine sponges."  
 RT Biol. Bull. 197:198-206(2000).  
 DR EMBL: Y17877; CAC14729.1; -.  
 DR HSP: P06213; IIRK.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002011; RTK\_kinaseII.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00069; pkinase.1.

DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR SMART; SM00220; S\_TKC; 1.  
 DR SMART; SM00219; TYKC; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
 DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
 DR ATP-binding; Kinase; transferase.  
 KW ATP-binding; Kinase; transferase; D57324138DBE1E4 CRC64;  
 SQ SEQUENCE 616 AA; 69477 MW; D57324138DBE1E4 CRC64;

Query Match 2.0%; Score 9; DB 5; Length 616;  
 Best Local Similarity 100.0%; Pred.No.2.2;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 SYGVVLMEM 199  
 DB 425 SYGVVLMEM 433

UNIT 24  
 AC 09A0J5; PRELIMINARY; PRT; 652 AA.  
 DT 01-JUN-2001 (TRENBLREL. 17, Created)  
 DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update)  
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)  
 DE DNA\_Ligase (EC 6.5.1.2) (Polydeoxyribonucleotide synthase [NAD+]).  
 GN GN OR SPY0751.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
 OC Streptococcaceae; Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SF370 / ATCC 700294 / SERO TYPE M1;  
 RC MEDLINE=21192684; PubMed=11296296;  
 RA Perrelet J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K., Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J., Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;  
 RA "Complete genome sequence of an M1 strain of Streptococcus pyogenes."  
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).  
 RL -1- FUNCTION: THIS PROTEIN CATALYZES THE FORMATION OF PHOSPHODIESTER LINKAGES BETWEEN 5'-PHOSPHORYL AND 3'-HYDROXYL GROUPS IN DOUBLE-STRANDED DNA USING NAD AS A COENZYME AND AS THE ENERGY SOURCE FOR THE REACTION. IT IS ESSENTIAL FOR DNA REPLICATION AND REPAIR OF DAMAGED DNA (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: NAD(+) + (DEOXYRIBONUCLEOTIDE) (N) + (DEOXYRIBONUCLEOTIDE) (M) = AMP + NICOTINAMIDE NUCLEOTIDE + (DEOXYRIBONUCLEOTIDE) (N+M).  
 CC -1- SIMILARITY: BELONGS TO THE NAD-DEPENDENT DNA LIGASE FAMILY.  
 CC EMBL: AE006527; AK33695.1; -.  
 DR HSP; O87703; 1B04.  
 DR InterPro; IPR001357; BRCT.  
 DR InterPro; IPR001679; DNALigase.  
 DR InterPro; IPR004150; DNA\_Ligase\_OB.  
 DR InterPro; IPR000445; HNH.  
 DR InterPro; IPR003583; HHH 1.  
 DR InterPro; IPR004149; Znf\_DNALigase\_C4.  
 DR Pfam; PF00533; BRCT; 1.  
 DR Pfam; PF01653; DNA\_Ligase\_N; 1.  
 DR Pfam; PF03120; DNA\_Ligase\_OB; 1.  
 DR Pfam; PF03119; DNA\_Ligase\_ZB; 1.  
 DR ProDom; PD003944; DNALigase; 1.  
 DR SMART; SM00292; BRCT; 1.  
 DR SMART; SM00278; HNH; 3.  
 DR SMART; SM00532; LIGANC; 1.  
 DR TIGRFAMs; TIGR00575; dnlj; 1.  
 DR PROSITE; PS50172; BRCT; 1.  
 DR PROSITE; PS01055; DNA\_Ligase\_N1; 1.  
 DR PROSITE; PS01056; DNA\_Ligase\_N2; 1.

KW DNA repair; DNA replication; Ligase; NAD; Complete Proteome.  
 SQ SEQUENCE 652 AA; 72420 MW; B88D93C9FD38B5 CRC64;

Query Match 2.0%; Score 9; DB 16; Length 652;  
 Best Local Similarity 100.0%; Pred.No.2.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 EKEAEILSV 59  
 DB 307 EKEAEILSV 315

RESULT 25

ID 012500; PRELIMINARY; PRT; 764 AA.  
 AC 012500;  
 DT 01-NOV-1996 (TRENBLREL. 01, Created)  
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)  
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)  
 DE Chromosome XII reading frame ORF YLR114C.  
 DE EFR4 OR L9354.8 OR L2941 OR YLR114C.  
 GN Saccharomyces cerevisiae (Baker's yeast).  
 OS Saccharomycetes; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Vernasek P., Voet M., Volckaert G.;  
 RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RA MIPS;  
 RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 RL [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C;  
 RA Geisel C.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RL [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C;  
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z., Favello A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D., Johnston L., Langston Y., Letreille P., Le T., Maris E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D., Rifken L., Riles L., Teich A., Trevisan E., Vignati D., Wilcox L., Wohldman P., Vaudin M., Wilson R., Waterson R.;  
 RA Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RL [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C;  
 RA Waterson R.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z73286; CA97681.1; -.  
 DR EMBL; U53878; AB867559.1; -.  
 DR EMBL; X89514; CA61692.1; -.  
 DR SCD; S0004104; EFR4.  
 SQ SEQUENCE 764 AA; 86425 MW; 159A00CF933A5C4A CRC64;

Query Match 2.0%; Score 9; DB 3; Length 764;  
 Best Local Similarity 100.0%; Pred.No.2.6;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGE 445  
 DB 628 DDDDDDDGE 636

RESULT 26  
 Q9XP8 PRELIMINARY; PRT; 789 AA.  
 AC Q9XP8;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Serine/threonine kinase PKN11.  
 GN PKN11.  
 OS Myxococcus xanthus.  
 OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;  
 OC Myxococcales; Cytochromatellaceae; Myxococcaceae; Myxococcus.  
 NCBI\_TaxID=34;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DFE1;  
 RA Inouye S., Jain R., Ueki T., Nariya H., Xu C., Hsu M.,  
 RA Munoz-Dorado J., Fares-Vidal E., Inouye M.,  
 RT "Sequence Analysis of 13 Eukaryotic-like Protein Ser/Thr Kinases of  
 RT Myxococcus xanthus, a Developmental Bacterium and Significance of  
 RT Their Coexistence with Protein His Kinases."  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF159690; AAD42854.1; -;  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00069; pkinase; 1.  
 DR ProDom: PD000001; Euk\_pkinase; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.  
 SQ SEQUENCE 789 AA; 81613 MW; EELE764F0BD77C63 CRC64;

Query Match 2.0%; Score 9; DB 2; Length 789;  
 Best Local Similarity 100.0%; Pred. No. 2.7;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 GVALWEMLT 201  
 Db 210 GVALWEMLT 218

## RESULT 27

Q9ZSM8 PRELIMINARY; PRT; 856 AA.

AC Q9ZSM8; 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE E2A1 (ATG402020/T10M13.3).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucotsid II; Brassicales; Brassicaceae; Arabidopsids.  
 NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Blodden P., Luo M., Dennis E.S., Peacock W.J., Chaudhury A.M.;  
 RT "E2A1, a novel polycarb group gene from Arabidopsis thaliana";  
 RL Plant Physiol. 0:0-0(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.

RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,  
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,  
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,  
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,  
 RA Ecker J.R.;  
 RT "Arabidopsis cDNA clones";  
 RN Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.

RA Kim C.J., Chen H., Cheuk R., Meyers M.C., Shinn P., Banh J.,  
 RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,  
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,

RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,  
 RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,  
 RA Ecker J.R.;  
 RT "Arabidopsis ORF clones";  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1 SIMILARITY: CONTAINS 1 SET DOMAIN.  
 DR EMBL: AF100163; AAD09108.1; -;  
 DR EMBL: AF057477; AAL09711.1; -;  
 DR EMBL: AY090293; AAL09554.1; -;  
 DR HSSP: P10969; 1WGT.  
 DR InterPro: IPR001005; Myb\_DNA\_binding.  
 DR InterPro: IPR001214; SET.  
 DR InterPro: IPR002221; WAP.  
 DR Pfam: PF00856; SET\_1.  
 DR SMART: SM00395; SANT; 1.  
 DR SMART: SM00317; SET; 1.  
 DR PROSITE: PS00317; 4 DISULFIDE\_CORE; UNKNOWN\_1.  
 DR PROSITE: PS00280; SET; 1.  
 SQ SEQUENCE 856 AA; 95396 MW; DD4B099C936F197C CRC64;

Query Match 2.0%; Score 9; DB 10; Length 856;  
 Best Local Similarity 100.0%; Pred. No. 2.9;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 DDDDDGSEE 447  
 Db 17 DDDDDGSEE 25

## RESULT 28

O04246 PRELIMINARY; PRT; 898 AA.

AC O04246; 01-JUN-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUN-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-MAR-1999 (TrEMBLrel. 20, Last annotation update)  
 DE Putative CURTLEAF-like 1 homeotic protein (POLYCOMB group-like  
 DE protein).  
 GN T10M13.3 OR ATG402020.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucotsid II; Brassicales; Brassicaceae; Arabidopsids.  
 NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Johnson A.F., de la Baeride M., Lohi M., Hoffman J., Hasegawa A.,  
 RA Gnoj L., Gottesman T., Grant S., Hamed A., Kaplan N., Schutz K.,  
 RA Shohdy N., Van Keuren K., Parnell L., Dedhia N., Martienssen R.,  
 RA McCombie M.;  
 RT "The sequence of the Arabidopsis thaliana T10M13 BAC";  
 RN Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.

RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Baeride M., Val D.M.,  
 RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,  
 RA Sheher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,  
 RA Mewes H.W., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1 SIMILARITY: CONTAINS 1 SET DOMAIN.  
 DR EMBL: AF001308; AAC78694.1; -;  
 DR EMBL: AL161493; CAB80695.1; -;  
 DR HSSP: P10969; 1WGT.  
 DR InterPro: IPR001005; Myb\_DNA\_binding.  
 DR InterPro: IPR001214; SET.  
 DR InterPro: IPR002221; WAP.  
 DR Pfam: PF00856; SET; 1.



RP	SEQUENCE FROM N.A.
RX	MEDLINE=94100173; PubMed=8274451;
RA	Schultz S.J., Nigg E.A.;
RT	"Identification of 21 novel human protein kinases, including 3 members of a family related to the cell cycle regulator nimA of Aspergillus nidulans.";
RL	Cell Growth Differ. 4:821-830(1993).
DR	EMBL; Z25430; CAA60917.1; -
DR	InterPro; IPR000719; Euk_kinase.
DR	InterPro; IPR01245; Tyr_kinase.
PFam	PF00069; kinase.1.
DR	PRINTS; PRD0109; TYRKINASE.
DR	PRODOM; PD000001; Euk_kinase.1.
DR	PROSITE; PSS0011; PROTEIN KINASE_DOM; 1.
KM	Kinase; Tyrosine-protein Kinase.
FT	NON_TER 1 101
FT	NON_TER 1 101
SEQ	SEQUENCE 101 AA; 11485 MW; E584845C18F4FCD8 CRC64;
 Query Match	
Best Local Similarity 100.0%; Score 8; DB 4; Length 101;	
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	172 WMAPEVIQ 179       
Db	92 WMAPEVIQ 99
 RESULT 34	
ID	Q9J872 PRELIMINARY; PRT; 106 AA.
AC	Q9J872;
DT	01-OCT-2000 (TREMBLrel. 15, Created)
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE	ORF3.
OS	Spodoptera exigua nucleopolyhedrovirus.
OC	Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC	Nucleopolyhedrovirus.
OX	NCBI_TaxID=10454;
RN	[1]
RN	SEQUENCE FROM N.A.
RP	MEDLINE=20036646; PubMed=10567663;
RX	Iutkel W.F., van Strien E.A., Heldens J.G., Broer R., Zuidema D.,
RA	Goldbach R.W., Vlak J.M.;
RT	"Sequence and organization of the spodoptera exigua multicapsid
RT	nucleopolyhedrovirus genome.";
RL	J. Gen. Virol. 80:3289-3304(1999).
RN	[2]
RP	SEQUENCE FROM N.A.
RA	Iutkel W.F., van Strien E.A., Heldens J.G.M., Broer R., Zuidema D.,
RA	Goldbach R.W., Vlak J.M.;
RL	Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF169823; AAF33593.1; -
SEQ	SEQUENCE 106 AA; 11780 MW; 599AA3E43F88023 CRC64;
 Query Match	
Best Local Similarity 100.0%; Score 8; DB 12; Length 106;	
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	436 GDDDDDDD 443       
Db	11 GDDDDDD 18
 RESULT 35	
ID	O55500 PRELIMINARY; PRT; 157 AA.
AC	O55500;
DT	01-JUN-1998 (TREMBLrel. 06, Created)
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE	UT51.

CN UL51.  
 OS Human cytomegalovirus (strain Towne).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Betaherpesvirinae; Cytomegalovirus.  
 OX NCBI\_TaxID=10363;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TOWNE.  
 RX MEDLINE=98241707; PubMed=9573236;  
 RA Krosky P.M., Underwood M.R., Turk S.R., Feng K.W., Jain R.K.,  
 RA Prak R.G., Western A.C., Birn K.K., Townsend L.B., Drach J.C.;  
 RT "Resistance of human cytomegalovirus to benzimidazole ribonucleosides  
 RT maps to two open reading frames: UL89 and UL56."  
 RL J. Virol. 72:4721-4728(1998).  
 DR EMBL: AF039234; AAC59474.1; -;  
 DR InterPro: IPR005208; Herpes\_UL33.  
 DR Pfam: PF03581; Herpes\_UL33; 1.  
 SQ SEQUENCE 157 AA; 16978 MW; 5999035AC484517D CRC64;

Query Match 1.8%; Score 8; DB 12; Length 157;  
 Best Local Similarity 100.0%; Pred. No. 7.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 440 DDDGDEE 447  
 Db 12 DDDGDEE 19

RESULT 36  
 Q9J867 PRELIMINARY; PRT; 161 AA.

AC Q9J867; 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE ORF68.  
 OS Spodoptera exigua nucleopolydnavirus.  
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
 OC Nucleopolydnavirus.  
 OX NCBI\_TaxID=10454;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20036646; PubMed=10567663;  
 RA Iukel W.F., van Strien E.A., Heidens J.G., Broer R., Zuidema D.,  
 RA Goldbach R.W., Vlak J.M.;  
 RT "Sequence and organization of the Spodoptera exigua multicapsid  
 RT nucleopolydnavirus genome."  
 RL J. Gen. Virol. 80:3289-3304(1999).  
 [2]

RA SEQUENCE FROM N.A.  
 RA Iukel W.F., van Strien E.A., Heidens J.G.M., Broer R., Zuidema D.,  
 RA Goldbach R.W., Vlak J.M.;  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF169823; AAF3598.1; -;  
 DR InterPro: IPR002557; Chitin\_bind\_Pera.  
 DR Pfam: PF01607; CEM\_14; 1.  
 DR SMART: SM00494; ChEBP2; 1.  
 SQ SEQUENCE 161 AA; 18509 MW; AE3BA84AD73A9B73 CRC64;

Query Match 1.8%; Score 8; DB 12; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 7.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
 Db 142 DDDDDDDG 149

RESULT 37  
 Q91902 PRELIMINARY; PRT; 161 AA.  
 AC Q91902;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE CUN017 hypothetical protein.  
 GN CUN017.  
 OS Culex nigripalpus baculovirus.  
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae.  
 OX NCBI\_TaxID=130556;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FLORIDA1997;  
 RX MEDLINE=2148685; PubMed=11602755;  
 RA Afonso C.L., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A.,  
 RA Becnel J.T., Rock D.L., Kutish G.F.;  
 RT "Genome Sequence of a Baculovirus Pathogenic for Culex nigripalpus."  
 RL J. Virol. 75:11157-11165(2001).  
 [2]

RN SEQUENCE FROM N.A.  
 RP STRAIN=FLORIDA1997;  
 RC Afonso C.L., Tulman E.R., Lu Z., Balinsky C.A., Moser B.A.,  
 RA Becnel J.T., Rock D.L., Kutish G.F.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF403738; AAK94095.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 161 AA; 18553 MW; 99BA93D3FB8386DD CRC64;

Query Match 1.8%; Score 8; DB 12; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 7.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
 Db 116 DDDDDDDG 123

RESULT 38  
 Q9HWQ2 PRELIMINARY; PRT; 162 AA.

AC Q9HWQ2; 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)  
 DE Hypothetical protein PA4129.  
 DE PA4129.

GN Pseudomonas aeruginosa.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PA01;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stever C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 RT opportunistic pathogen."  
 RL Nature 406:959-964(2000).  
 DR EMBL: AE004829; AAG07516.1; -;  
 KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 162 AA; 18122 MW; 094CA42A8687795 CRC64;

Query Match 1.8%; Score 8; DB 16; Length 162;  
 Best Local Similarity 100.0%; Pred. No. 7.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 PULPLAA 336  
 Db 34 PULPLAA 41



```

RESULT 39
SQ 09840 PRELIMINARY; PRT; 180 AA.
AC 09840;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE Essential YAEI protein.
GN YAEI.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Microsporici Saccharomycetales; Candida.
OX NCBI_Taxid=5476;
RN [1]
RP SEQUENCE FROM N.A.
RA De Backer M.D., Logghe M., Vlaene J., Loonen I., Vandoninck S.,
RA Contreiras R., Luyten W.H.M.L.;
RA "A novel method for systematic identification of genes required for
growth of Candida albicans."
Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
EMBL: AJ190499; CAB7639.1;
SQ SEQUENCE 180 AA; 20240 MW; EC98B8E63C1676C7 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 3; Length 180;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 436 GDDDDDD 443
Db 117 GDDDDDD 124

RESULT 40
SQ 000193 PRELIMINARY; PRT; 183 AA.
AC 000193;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DE Small acidic protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Carnivora; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=PLACENTA;
RA MEDLINE=97408516; Pubmed=9263035;
RA Gong T.W., Hegeman A.D., Shin J.J., Lindberg K.H., Baraid K.F.,
RA Lomax N.I.;
RA "Novel genes expressed in the chick oocyte during development:
identification using differential display of RNA."
Int. J. Dev. Neurosci. 15:585-594 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RA TISSUE=BRIN;
RA Strauberg R.;
RA Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA TISSUE=BRIN;
RA Strauberg R.;
RA Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA TISSUE=BRIN;
RA Strauberg R.;
RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
EMBL: U51678; AAB98729.1;
RA EMBL: BC016352; AAH16352.1;
RA EMBL: BC007103; AAH07103.1;
RA EMBL: BC020937; AAH20937.1;

```

```

SQ SEQUENCE 183 AA; 20332 MW; 509BA7377B699A74 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 4; Length 183;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 436 GDDDDDD 443
Db 118 GDDDDDD 125

RESULT 41
SQ 042341 PRELIMINARY; PRT; 184 AA.
AC 042341;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE Serine-threonine protein kinase (Fragment).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_Taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=CELL SUSPENSION CULTURE OF ECOTYPE COLUMBIA;
RA Cooke R., Laudie M., Raynal M., Delzeny M.;
RA Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
EMBL: F1996; CA23375.1;
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase.1.
DR PRINTS: PRO0109; TYRKINASE.
DR PRODOM: PD000001; Euk_pkinase.1.
DR PROSITE: PS0011; PROTEIN KINASE_DOM; 1.
FT KINASE; Tyrosine-protein Kinase.
FT NON TER 1
FT NON TER 184
SQ SEQUENCE 184 AA; 20933 MW; EB93761356ACE2E0 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 10; Length 184;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 172 WMAPEVIO 179
Db 137 WMAPEVIO 144

RESULT 42
SQ 090400 PRELIMINARY; PRT; 194 AA.
AC 090400;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE Hypothetical 22.3 kDa protein.
OS White spot syndrome virus (WSSV).
OC Viruses; unclassified viruses.
OX NCBI_Taxid=92652;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20112891; Pubmed=10644828;
RA van Hulten M.C., Teal M.F., Schipper C.A., Lo C.F., Kou G.H.,
RA Vlak J.M.;
RA "Analysis of a genomic segment of white spot syndrome virus of shrimp
containing ribonuclease reductase genes and repeat regions."
J. Gen. Virol. 81:307-316 (2000).
EMBL: AF099142; AAF18485.1;
KM Hypothetical protein.
SQ SEQUENCE 194 AA; 22267 MW; 73254CF17FE9678C CRC64;

```

Query Match 1.8%; Score 8; DB 12; Length 194;  
 Best Local Similarity 100.0%; Pred. No. 8.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
 61 GDDDDDD 68

DB 61 GDDDDDD 68

RESULT 43

Q91LE7 PRELIMINARY; PRT; 194 AA.

AC Q91LE7; 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE ORP94.  
 OS White spot syndrome virus (WSSV).  
 OC Viruses; unclassified viruses.  
 NCBI\_TaxID=92652;

[1]  
 SEQUENCE FROM N.A.  
 MEDLINE=21342572; PubMed=11448154;  
 RA van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,  
 RA Tarchini R., Fiers M., Sandbrink H., Lankhorst R.K., Vlak J.M.;  
 RL "The white spot syndrome virus DNA genome sequence.";  
 RT Virology 286:7-22(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,  
 RA Tarchini R., Fiers M., Sandbrink H., Lankhorst R.K., Vlak J.M.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF369029; AAK77763.1; -  
 SQ SEQUENCE 194 AA; 22281 MW; 19FE26A0504DC78 CRC64;

Query Match 1.8%; Score 8; DB 12; Length 194;  
 Best Local Similarity 100.0%; Pred. No. 8.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
 61 GDDDDDD 68

DB 61 GDDDDDD 68

RESULT 44

Q90TE0 PRELIMINARY; PRT; 194 AA.

AC Q90TE0; 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE WSSV234;  
 OS White spot syndrome virus (WSSV).  
 OC Viruses; unclassified viruses.  
 NCBI\_TaxID=92652;

[1]  
 SEQUENCE FROM N.A.  
 MEDLINE=20517548; PubMed=11062040;  
 RA Tsai M.F., Yu H.T., Tzeng H.F., Lue J.H., Chou C.M., Huang C.J.,  
 RA Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;  
 RL "Identification and characterization of a shrimp white spot syndrome  
 virus (WSSV) gene that encodes a novel chimeric polypeptide of  
 cellular-type thymidine kinase and thymidylate kinase.";  
 RT Virology 277:100-110(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-TAIWAN;  
 RC MEDLINE=21844071; PubMed=11853398;  
 RA Chen L.L., Lue J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,  
 RA Lo C.F., Kou G.H.;  
 RL "Identification of a nucleocapsid protein (VP35) gene of shrimp white  
 spot syndrome virus and characterization of the motif important for

targeting VP35 to the nuclei of transfected insect cells.";  
 RL Virology 293:44-53(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TAIWAN;  
 RA Lo C.-F., Kou G.-H.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF440570; AAL89102.1; - 73FFC0B214CCD3D CRC64;  
 SQ SEQUENCE 194 AA; 22253 MW; 73FFC0B214CCD3D CRC64;

Query Match 1.8%; Score 8; DB 12; Length 194;  
 Best Local Similarity 100.0%; Pred. No. 8.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
 61 GDDDDDD 68

DB 61 GDDDDDD 68

RESULT 45

Q9CTY9 PRELIMINARY; PRT; 198 AA.

AC Q9CTY9; 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Hypothetical 21.8 kDa protein (Unknown protein).  
 GN T2015.12.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.

[1]  
 SEQUENCE FROM N.A.  
 RC STRAIN=CY. COLUMBIA;  
 RX MEDLINE=2106719; PubMed=11130712;  
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,  
 RA Brehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
 RA Dunn P., Egn P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gail J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B.,  
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Malet R., Marzilli A.,  
 RA Miltscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 RA Sun H., Tallon L.U., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Utechtback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RL "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 thaliana.";  
 RT Nature 408:816-820(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,  
 RA Lam B., Southwick A., Jones T., Banu J., Carlini P., Chen H.,  
 RA Palm C.J., Bowser L., Jones T., Banu J., Carlini P., Chen H.,  
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Sakano H.,  
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
 RA Ecker J., Theologis A., Davis R.W.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC051631; AAG51532.1; -  
 DR EMBL; AF386997; AAK62442.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 198 AA; 21847 MW; COAC344A0FE7818 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 8.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
DB 101 GDDDDDD 108

## RESULT 46

ID 039756 PRELIMINARY; PRT; 204 AA.

AC 039756;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE PK1 protein (Fragment).  
GN PK1.  
OS Fagus sylvatica (Beechnut).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fagales; Fagaceae; Fagus.  
NCBI\_TaxID=28930;  
[1]

SEQUENCE FROM N.A.  
RA Nicolas C., Nicolas G., Rodriguez D.;  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; X97547; CA66149.1; -.  
DR HSP; P08631; IAD5.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW Serine/threonine-protein kinase.  
FT NON\_TER  
SQ SEQUENCE 204 AA; 23305 MW; B4C70D228B1077B0 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 204;  
Best Local Similarity 100.0%; Pred. No. 9.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VIHRLDKS 136  
DB 22 VIHRLDKS 29

## RESULT 47

ID 09SWT7 PRELIMINARY; PRT; 206 AA.

AC 09SWT7;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE HVI.PST.  
GN HVI.PST.  
OS Hordeum vulgare (Barley).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;  
OC Triticeae; Hordeum.  
NCBI\_TaxID=4513;  
[1]

SEQUENCE FROM N.A.  
RA MEDLINE=99324224; PubMed=10393983;  
RA Feillet C., Keller B.;  
RT "High gene density is conserved at syntenic loci of small and large  
RT grass genomes."  
RL Proc. Natl. Acad. Sci. U.S.A. 96:8265-8270 (1999).  
RA EMBL; AF108009; AAD6470.1;  
SQ SEQUENCE 206 AA; 21440 MW; 3F0BF6622035942 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 206;

Best Local Similarity 100.0%; Pred. No. 9.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 436 GDDDDDD 443  
DB 184 GDDDDDD 191

## RESULT 48

ID 0945P0 PRELIMINARY; PRT; 209 AA.

AC 0945P0;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Transcriptional activator FHA1.  
GN FHA1.  
OS Nicotiana tabacum (Common tobacco).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
NCBI\_TaxID=4097;  
[1]

SEQUENCE FROM N.A.  
RA STRAIN=CV. XANT1;  
RA Kim M.I., An J.W., Pai H.S.;  
RT "FHA1, transcription activator containing forkhead-associated  
RT domain."  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF11856; AAL05884.1; -.  
DR InterPro; IPR000253; FHA\_domain.  
DR Pfam; PF00498; FHA; 1.  
DR PROSITE; PS50006; FHA\_DOMAIN; 1.  
SQ SEQUENCE 209 AA; 22601 MW; E3440C7327070176 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 209;  
Best Local Similarity 100.0%; Pred. No. 9.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
DB 164 DDDDDDDG 171

## RESULT 49

ID 028960 PRELIMINARY; PRT; 211 AA.

AC 028960;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Hypothetical protein AF1309.  
GN AF1309.  
OS Archaeoglobus fulgidus.  
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
OC Archaeoglobaceae; Archaeoglobus.  
NCBI\_TaxID=2234;  
[1]

SEQUENCE FROM N.A.  
RA STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
RA MEDLINE=98049343; PubMed=9389475;  
RA Klein H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,  
RA Fleischmann D.L., Kerlavage A.R., Graham D.B., Kyriades N.C.,  
RA Kirkness S.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
RA Cotton M.D., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,  
RA Sadow P.W., Spriggs K.P., Bowman C., Kalne B.P., Sykes S.M.,  
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
RA Venter J.C.;  
RT "The complete genome sequence of the hyperthermophilic, sulphate-

RT reducing archaeon *Archaeoglobus fulgidus*.  
 RL Nature 390:364-370(1997).  
 DR EMBL; AF001013; AAB8938.1; -  
 DR TIGR; AF1309; -  
 DR InterPro; IPR002741; DUF56.  
 DR Pfam; PF01879; DUF56; 1.  
 DR Hypothetical protein; Complete proteome.  
 KW SEQUENCE 211 AA; 2282 MW; EB258FASD12855CA CRC64;  
 SQ

Query Match 1.8%; Score 8; DB 17; Length 211;  
 Best Local Similarity 100.0%; Pred. No. 9.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 DDDDDDDG 336  
 DB 174 PLLEPLAA 181

## RESULT 50

Q91VH4 PRELIMINARY; PRT; 213 AA.  
 Q91VH4  
 01-DEC-2001 (TRENBLrel. 19, Created).  
 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
 DE Hypothetical 23.6 kDa protein (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRST TUMOR;  
 RA Strausberg R.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC014284; AAH14284.1; -  
 DR InterPro; IPR002038; Osteopontin.  
 DR Pfam; PF00865; Osteopontin; 1.  
 KW Hypothetical protein.  
 FT  
 NON\_TER 1  
 SQ SEQUENCE 213 AA; 23625 MW; 50CCE5DDE9E91797 CRC64;

Query Match 1.8%; Score 8; DB 11; Length 213;  
 Best Local Similarity 100.0%; Pred. No. 9.5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
 7 DDDDDDDG 14

## RESULT 51

Q8RZR9 PRELIMINARY; PRT; 218 AA.  
 Q8RZR9  
 01-JUN-2002 (TRENBLrel. 21, Created)  
 01-JUN-2002 (TRENBLrel. 21, Last sequence update)  
 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
 DE B1140D12.10 protein.  
 GN B1140D12.10  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Eukaryota; Viridiplantae; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eurythaleales; Oryzaceae; Oryza.  
 NCBI\_Taxid=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T.; Matsumoto T.; Yamamoto K.;  
 RT "Oryza sativa nipponbare (GAS) genomic DNA, chromosome 1, BAC  
 clone: B1140D12."  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP003537; BAB86162.1; -

SQ SEQUENCE 218 AA; 23468 MW; 9B158986E738F29 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 218;  
 Best Local Similarity 100.0%; Pred. No. 9.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
 DB 204 DDDDDDDG 211

## RESULT 52

Q8X4P6 PRELIMINARY; PRT; 229 AA.  
 Q8X4P6  
 01-MAR-2002 (TRENBLrel. 20, Created)  
 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
 DE Putative antiterminal Q of prophage CP-933X (Antitermination  
 protein).  
 GN Z1874 OR EC31620  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 NCBI\_Taxid=83334;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T.; Plunkett G. III; Burland V.; Mau B.; Glaesner J.D.;  
 Rose D.J.; Mayhew G.F.; Evans P.S.; Gregor J.; Kirkpatrick H.A.;  
 Posfai G.; Hackett J.; Klink S.; Boutin A.; Shao Y.; Miller L.;  
 Grobbeck E.J.; Davis N.W.; Lim A.; Dimalanta E.T.; Potamoukis K.;  
 Apodaca J.; Anantharaman T.S.; Lin J.; Yen G.; Schwartz D.C.;  
 Welch R.A.; Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7."  
 RL Nature 409:529-533(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RIMD 0509952;  
 MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T.; Makino K.; Onishi M.; Kurokawa K.; Ishii K.; Yokoyama K.;  
 Han C.-G.; Ohtsubo E.; Nakayama K.; Murata T.; Tanaka M.; Tobe T.;  
 Iida T.; Takami H.; Honda T.; Sasaki K.; Ogasawara N.; Yasunaga T.;  
 RA Kubara S.; Shiba T.; Hattori M.; Shingawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*  
 O157:H7 and genomic comparison with a laboratory strain K-12."  
 RL DNA Res. 8:11-22(2001).  
 DR EMBL; AF005330; AAG55971.1; -  
 DR EMBL; AP002555; BAB35043.1; -  
 KW Complete proteome.  
 SQ SEQUENCE 229 AA; 26402 MW; E6A8F81BE9476667 CRC64;

Query Match 1.8%; Score 8; DB 16; Length 229;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 290 TIERIKKL 297  
 DB 142 TIERIKKL 149

## RESULT 53

Q50439 PRELIMINARY; PRT; 237 AA.  
 Q50439  
 01-NOV-1996 (TRENBLrel. 01, Created)  
 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
 DE Tmp protein.  
 GN TmpR.  
 OS Mycobacterium smegmatis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1772;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92078082; PubMed=1660454;  
 RA Cutillo J.D., Barletta R.G., Bloom B.R., Jacobs W.R. Jr.;  
 RT "A novel transposon trap for mycobacteria: isolation and  
 RL characterization of IS1096."  
 DR J. Bacteriol. 173:7772-7780(1991).  
 EMBL: M76495; AAA98488.1;  
 SQ SEQUENCE 237 AA; 26293 MW; B4A0C4952D5C68BE CRC64;

Query Match 1.8%; Score 8; DB 2; Length 237;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 431 FDLSEGGD 438  
 Db 119 FDLSEGGD 126

## RESULT 54

O9AS90 PRELIMINARY; PRT; 246 AA.  
 AC O9AS90;  
 DT 01-JUN-2001 (T-EMBLrel. 17, Created)  
 DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)  
 DE P0028E10.1 protein (P0701D05.18 protein).  
 GN P0028E10.1 OR P0701D05.18.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nippobare(GM3) genomic DNA, chromosome 1, PAC  
 RL clone: P0028E10."  
 RN Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nippobare(GM3) genomic DNA, chromosome 1, PAC  
 RL clone: P0701D05."  
 RN Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP003301; BAB64804.1;  
 SQ SEQUENCE 246 AA; 25160 MW; F9B39125F8A5A534 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 246;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDD 443  
 Db 179 GDDDDDD 186

## RESULT 55

O9M1H5 PRELIMINARY; PRT; 250 AA.  
 AC O9M1H5;  
 DT 01-OCT-2000 (T-EMBLrel. 15, Created)  
 DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)  
 DE 01-OCT-2000 (T-EMBLrel. 15, Last annotation update)  
 GN Hypothetical 29.2 kDa protein.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX Obermayer B., Ottenwaelder B., Duchemin D., Zeitler K., Mewes H.W.,  
 RA Rudd S., Lemke K., Mayer K.F.X., Quetier F., Salanoubat M.;  
 RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL138656; CAB7572.1;  
 SQ SEQUENCE 250 AA; 29183 MW; E153440480B8B5F6 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 RLKRLERD 300  
 Db 54 RLKRLERD 61

## RESULT 56

O9GR09 PRELIMINARY; PRT; 258 AA.  
 AC O9GR09;  
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)  
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)  
 DE 01-MAR-2001 (T-EMBLrel. 16, Last annotation update)  
 GN L6520.05.  
 OS Leishmania major.  
 OC Eukaryota; Euzoenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
 OX NCBI\_TaxID=5664;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FRIEDLIN;  
 RA Masny D., Burnelle B., Goffeau A., Ivens A.C., Quail M.,  
 RA Rajandream M.A., Barrell B.G.;  
 RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FRIEDLIN;  
 RX MEDLINE=98146435; PubMed=9477341;  
 RA Ivens A.C., Lewis S.W., Bagherzadeh A., Zhang L., Chan H.M.,  
 RA Smith D.F.;  
 RT "A physical map of the Leishmania major Friedlin genome."  
 RL Genome Res. 8:135-145(1998).  
 DR EMBL: AL446005; CAC14532.1;  
 KW Hypothetical protein.  
 SQ SEQUENCE 258 AA; 28116 MW; 52A1CB8119F2B079 CRC64;

Query Match 1.8%; Score 8; DB 5; Length 258;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 436 GDDDDDD 443  
 Db 70 GDDDDDD 77

## RESULT 57

O9CXN1 PRELIMINARY; PRT; 259 AA.  
 AC O9CXN1;  
 DT 01-JUN-2001 (T-EMBLrel. 17, Created)  
 DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)  
 DE 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)  
 GN 3110052N05RIK protein.  
 GN 3110052N05RIK.

```

OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=EMBRYONIC HEAD;
RX MEDLINE=2108566; PubMed=1127851;
RA Kawai J., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aikawa K., Izawa Y., Nishi K., Kiyosawa H., Kondo S., Yamatake I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasubawa T., Saito K.,
RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochava H.,
RA Knehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiraldi L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Maehio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barz G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guelincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Rodriguez I., Sakamoto N.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seyer T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kontecki S.,
RA Hayashizaki Y.;
RA "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
RL EMBL, AK014208; BAB39206.1; -.
DR MGI:1924237; 3110052N05R1.
DR InterPro: IPR001454; Hlgnaase/hydrilase.
DR Pfam: PF00702; Hydrilase; 1.
SQ SEQUENCE 259 AA; 28730 MW; C8B34D5FD6C9A8 CRC64;

Query Match 1.8%; Score 8; DB 11; Length 259;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 291 LERUKLE 298
DB 57 LERUKLE 64

RESULT 58
Q9S180 PRELIMINARY; PRT; 260 AA.
ID Q9S180
AC Q9S180;
RT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
GN At2g25220 protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotsids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.T., Mason T.M., Bowman C.L., Barnstead M.E., Feldguy T.V.,
RA Beil C.R., Ketchum K.A., Lee J.-J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.B., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana.";
RT Nature 402:761-768(1999).
[2]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X.;
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1-SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AC007070; AAD23669.1; -.
DR InterPro: IPR000719; Euk_Pkinase.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Euk_Pkinase; 1.
DR PROSITE: PS00107; PROTEIN KINASE ATP; UNKNOWN_1.
DR PROSITE: PS0011; PROTEIN KINASE DOM; 1.
DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 260 AA; 28793 MW; 5A469E1CB1596BEC CRC64;

Query Match 1.8%; Score 8; DB 10; Length 260;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VTHRDLS 136
DB 125 VTHRDLS 132

RESULT 59
Q948F2 PRELIMINARY; PRT; 277 AA.
ID Q948F2
AC Q948F2;
RT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE Putative polyprotein.
GN OSNBA0013J21.1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
[1]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
RA Saki C., Henry D., Oates R., Simmons J.;
RT "Rice Genomic Sequence";
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC079843; AAL01161.1; -.
DR POLYPROTEIN.
SQ SEQUENCE 277 AA; 31970 MW; 29730C9E1C3F2027 CRC64;

Query Match 1.8%; Score 8; DB 10; Length 277;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
DB 75 DDDDDDDG 82

RESULT 60
Q64125 PRELIMINARY; PRT; 279 AA.
ID Q64125
AC Q64125;
RT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE Early T-lymphocyte activator-1 (Fragment).
GN SPPI OR ETA-1/OP.
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
[1]
RP SEQUENCE FROM N.A.

```

RX MEDLINE=95272553; PubMed=7753053;  
 RA Ono M., Yamamoto T., Nose M.,  
 RT "Allelic difference in the nucleotide sequence of the Era-1/Op gene  
 transcritp.";  
 RL Mol. Immunol. 32:447-448 (1995).  
 DR EMBL: S78177; AAB3451.2; -.  
 DR MGD: MGI:98389; Sgpl.  
 DR InterPro: IPR02038; Osteopontin.  
 DR Pfam: PF00865; Osteopontin.1.  
 DR PRINTS: PR00216; OSTEOPTNTN.  
 DR SMART: SM00017; OSTEO.1.  
 DR PROSITE: PS00884; OSTEOPTNTN.1.  
 FT NON TER  
 SQ SEQUENCE 279 AA; 30939 MW; BOEF855D0AC14A2 CRC64;

Query Match 1.8%; Score 8; DB 11; Length 279;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 437 DDDDDGG 444  
 73 DDDDDGG 80

RESULT 61  
 002445 PRELIMINARY; PRT; 292 AA.  
 AC 002445;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE LOX6 (Fragment).  
 GN LOX6.  
 OS Helobdella robusta (Leech).  
 OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinea; Hirudinea;  
 OC Rhynchobdellida; Glossiphoniidae; Helobdella.  
 OX NCBI\_TaxID=6412;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98008903; PubMed=9344545;  
 RA Kourakis M.J., Master V.A., Lohrhorst D.K., Nardelli-Haeffliger D.,  
 RT Weiden C.J., Martindale M.Q., Shankland M.,  
 RT "Conserved anterior boundaries of Hox gene expression in the central  
 nervous system of the leech Helobdella."  
 RL Dev. Biol. 190:284-300 (1997).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL: AF004386; AAB61441.1; -.  
 DR HSSP: P02833; 9ANT.  
 DR TRANSFAC: T03754; -.  
 DR InterPro: IPR001827; Antennapedia.  
 DR InterPro: IPR001356; Homeobox.  
 DR Pfam: PF000047; HTH repressor.  
 DR Pfam: PF00046; homeobox.1.  
 DR PRINTS: PR00025; ANTENNAPEDIA.  
 DR PRINTS: PR00024; HOMEBOX.  
 DR Prodom: PD000010; Homeobox.1.  
 DR SMART: SM00389; HOX.1.  
 DR PROSITE: PS00032; ANTENNAPEDIA.1.  
 DR PROSITE: PS00027; HOMEBOX.1; 1.  
 DR PROSITE: PS00071; HOMEBOX.2; 1.  
 KW DNA-binding; Homeobox; Nuclear protein.  
 FT NON TER  
 SQ SEQUENCE 292 AA; 31973 MW; 66F82BBA65B5D674 CRC64;

Query Match 1.8%; Score 8; DB 5; Length 292;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 438 DDDDDGG 445  
 Db 194 DDDDDGG 201

RESULT 62  
 09NMX4 PRELIMINARY; PRT; 293 AA.  
 AC 09NMX4;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE CDNA FLJ20549 f18, clone KAT11561.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,  
 RA Suzuki Y., Ohtsushi M., Nishi T., Shibahara T., Tanaka T.,  
 RA Nakamura Y., Isega T., Sugano S.,  
 RT "NEDO human cDNA sequencing project.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK000556; BAA91251.1; -.  
 DR HSSP: Q00534; 1B18.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00069; Pkinase.1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR Prodom: PD000001; Euk\_pkinase.1.  
 DR PROSITE: PS00011; PROTEIN KINASE DOM.1.  
 KW ATP-binding; Transferase; Tyrosine-protein kinase.  
 SQ SEQUENCE 293 AA; 32883 MW; 49A1AB8BC97CB03E CRC64;

Query Match 1.8%; Score 8; DB 4; Length 293;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 172 WMAPEVQ 179  
 Db 62 WMAPEVQ 69

RESULT 63  
 017319 PRELIMINARY; PRT; 297 AA.  
 AC 017319;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE NK-2 class homeodomain protein.  
 GN CEH-24.  
 OS Caenorhabditis briggsae.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6238;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Harfe B., Fire A.,  
 RT Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL: AF026057; AAB86603.1; -.  
 DR HSSP: P23441; 1FTT.  
 DR TRANSFAC: T03376; -.  
 DR InterPro: IPR001356; Homeobox.  
 DR Pfam: PF00046; homeobox.1.  
 DR PRINTS: PR00024; HOMEBOX.  
 DR Prodom: PD000010; Homeobox.1.  
 DR SMART: SM00389; HOX.1.  
 DR PROSITE: PS00027; HOMEBOX.1; 1.  
 DR PROSITE: PS00071; HOMEBOX.2; 1.  
 KW DNA-binding; Homeobox; Nuclear protein.  
 SQ SEQUENCE 297 AA; 32886 MW; 3FAF5A4886D60755 CRC64;

Query Match 1.8%; Score 8; DB 5; Length 297;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
Db 224 GDDDDDD 231

## RESULT 64

064936 \* PRELIMINARY; PRT; 298 AA.  
AC 064936;  
DT 01-AUG-1998 (TREMBlrel. 07, Created)  
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Zinc finger protein (Fragment).  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
OC NCBI\_TaxID=3702;

## [1] SEQUENCE FROM N.A.

STRAIN=LADSEBEG ERECTA;  
RA Blodieu P., Luo M., Dennis E.S., Peacock W.J., Chaudhury A.M.;  
RT "Cloning of a cDNA encoding a two fingered C2H2 zinc-finger protein  
from Arabidopsis thaliana (Accession No. AF030304) (PGR98-048).";  
RL Plant Physiol. 116:1193-1193 (1998).  
DR EMBL; AF030304; AAC09174.1;  
DR InterPro; IPR000822; Znf\_C2H2.  
DR Pfam; PF00096; Znf\_C2H2\_2.  
DR SMART; SM00355; Znf\_C2H2\_2.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 2.  
DR PROSITE; PS00157; ZINC\_FINGER\_C2H2\_2; 2.  
KM DNA-binding; Metal-binding; zinc-finger.  
FT NON\_TER 1  
SEQUENCE 298 AA; 32954 MW; C3E290E9F4227A2D CRC64;

## Query Match

Best Local Similarity 1.8%; Score 8; DB 10; Length 298;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
Db 25 DDDDDDDG 32

## RESULT 65

08VB23 PRELIMINARY; PRT; 302 AA.

08VB23;  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE May178.  
OS White spot syndrome virus (WSSV).  
OC Viruses; unclassified viruses.  
OC NCBI\_TaxID=92652;

## [1] SEQUENCE FROM N.A.

RA MEDLINE=21548311; PubMed=11689662;  
RX Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;  
RT "Complete genome sequence of the shrimp white spot bacilliform  
virus.";  
RL J. Virol. 75:11811-11820 (2001).

## [2] SEQUENCE FROM N.A.

RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;  
RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF332093; AAL33182.1;  
SEQUENCE 302 AA; 35369 MW; 23C439039EA434A1 CRC64;

## Query Match

Best Local Similarity 1.8%; Score 8; DB 12; Length 302;  
Best Local Similarity 100.0%; Pred. No. 13;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
Db 61 GDDDDDD 68

## RESULT 66

09NKD3 PRELIMINARY; PRT; 307 AA.  
AC 09NKD3;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE Hypothetical 33.9 kDa protein.  
GN BG.DS05899.6.  
OS Drosophila melanogaster (fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Eukaryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidae; Drosophilidae; Drosophila.  
OC NCBI\_TaxID=7227;

## [1] SEQUENCE FROM N.A.

STRAIN=BERKELEY;  
RC MEDLINE=99403001; PubMed=10471707;  
RA Ashburner M., Miera S., Roote J., Lewis S.E., Blazej R., Davis T.,  
RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,  
RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,  
RA Palazolo N., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,  
RA Celinkner S., Rubin G.M.;  
RT "An exploration of the sequence of a 2.9-Mb region of the genome of  
Drosophila melanogaster: the Adh region.";  
RL Genetics 153:179-219 (1999).

## [2] SEQUENCE FROM N.A.

STRAIN=BERKELEY;  
RC Celinkner S.E., Agbayani A., Arcana T.T., Baxter E., Blazej R.G.,  
RA Butenloft C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,  
RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,  
RA Houston K.A., Hummasti S.R., Kaira K., Kearney L., Kim E., Lee B.,  
RA Lewis S., Li P., Lomocan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,  
RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,  
RA Sethi H., Smit E., Svyrskas R.R., Wan K.H., Weinburg T., Zhang R.,  
RA Zieran L.L., Rubin G.M.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE003408; AAP4852.1;  
DR FlyBase; FBgn0028873; BG.DS05899.6.  
KM Hypothetical protein.  
SEQUENCE 307 AA; 33948 MW; 24E40352F49F8C2F CRC64;

## Query Match

Best Local Similarity 1.8%; Score 8; DB 5; Length 307;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
Db 82 DDDDDDDG 89

## RESULT 67

P97827 PRELIMINARY; PRT; 317 AA.

AC P97827;  
DT 01-MAY-1997 (TREMBlrel. 03, Created)  
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Osteopontin.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OC NCBI\_TaxID=10116;  
SEQUENCE FROM N.A.



RC STRAIN-WISTER;  
 RA Yamamoto S.;  
 RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-WISTER;  
 RA YAMAMOTO S., UEMURA T.;  
 RT "Localization of osteopontin fused to the green fluorescent protein in  
 living osteoblastic cells."  
 RL Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AB001382; BAA19247.1; -  
 DR InterPro; IPR002038; Osteopontin.  
 DR Pfam; PF00865; Osteopontin; 1.  
 DR PRINTS; PR00216; OSTEOPONTIN.  
 DR SMART; SM00017; OSTEO; 1.  
 DR PROSITE; PS00884; OSTEOPONTIN; 1.  
 SQ SEQUENCE 317 AA; 34963 MW; 779A5F00F86E2310 CRC64;

Query Match 1.8%; Score 8; DB 11; Length 317;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 437 DDDDDDD 444  
 Db 88 DDDDDDD 95

RESULT 68  
 ID 09PTX7 PRELIMINARY; PRT; 318 AA.  
 AC 09PTX7;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Calreticulin (Fragment).  
 OS Lethenteron reissneri.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;  
 OC Petromyzontiformes; Petromyzontidae; Lethenteron.  
 OX NCBI\_TaxID=7753;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20063780; PubMed=10594174;  
 RA Kuraku S., Hoehiyama D., Katoh K., Suga H., Miyata T.;  
 RT "Monophyly of lampreys and hagfishes supported by nuclear DNA-coded  
 genes."  
 RL J. Mol. Evol. 49:729-735 (1999).  
 DR EMBL; AB025328; BAA88481.1; -  
 DR InterPro; IPR001580; Calreticulin.  
 DR InterPro; IPR000886; ER target.  
 DR Pfam; PF00262; calreticulin; 1.  
 DR PRINTS; PR00626; CALRETICULIN; 1.  
 DR PRODOM; PD001866; CALRETICULIN; 1.  
 DR PROSITE; PS00804; CALRETICULIN 2; 1.  
 DR PROSITE; PS00805; CALRETICULIN REPEAT; 3.  
 DR PROSITE; PS00014; ER TARGET; UNKNOWN\_1.  
 FT NON TER 1  
 SQ SEQUENCE 318 AA; 36997 MW; C88102EAC1C1506 CRC64;

Query Match 1.8%; Score 8; DB 13; Length 318;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 436 GDDDDDD 443  
 Db 285 GDDDDDD 292

RESULT 69  
 ID 027264 PRELIMINARY; PRT; 318 AA.  
 AC 027264;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Conserved protein.  
 GN MTH1196.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
 OC Methanobacteriaceae; Methanothermobacter.  
 OX NCBI\_TaxID=187420;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DELTA H;  
 RX MEDLINE=98037514; PubMed=9371463;  
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
 RA Altredge T., Baehrtraden R., Blakely D., Cook R., Gilbert K.,  
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pochler B., Qiu D.,  
 RA Spadafora R., Viçcare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jivani N., Caruso A., Bush D., Safer H., Patwell J., Prabhakar S.,  
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 strain delta: functional analysis and comparative genomics."  
 RL J. Bacteriol. 179:7135-7155 (1997).  
 DR EMBL; AE000888; AAB5685.1; -  
 DR InterPro; IPR003801; DUF198.  
 DR Pfam; PF02649; DUF198; 1.  
 DR Complete proteome.  
 KW Complete proteome.  
 SQ SEQUENCE 318 AA; 36035 MW; 75BC399BCD1535D9 CRC64;

Query Match 1.8%; Score 8; DB 17; Length 318;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 44 VKKLKIE 51  
 Db 28 VKKLKIE 35

RESULT 70  
 ID 0913E5 PRELIMINARY; PRT; 321 AA.  
 AC 0913E5;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE ORF1771 (Fragment).  
 OS White spot syndrome virus (WSSV).  
 OC Viruses; unclassified viruses.  
 OX NCBI\_TaxID=92652;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Zhang X., Huang C., Hew C.L.;  
 RT "A gene from shrimp white spot syndrome virus."  
 RL Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AY050248; AAI11022.1; -  
 FT NON TER 321  
 SQ SEQUENCE 321 AA; 35834 MW; 9D9083FEA1BE3E08 CRC64;

Query Match 1.8%; Score 8; DB 12; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 436 GDDDDDD 443  
 Db 170 GDDDDDD 177

RESULT 71  
 ID 08VAL2 PRELIMINARY; PRT; 321 AA.  
 AC 08VAL2;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Wav390.  
OS white spot syndrome virus (WSSV).  
OC Viruses; unclassified viruses.  
OX NCBI\_Taxid=92652;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21548311; PubMed=11689662;  
RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;  
RT "Complete genome sequence of the shrimp white spot bacilliform  
RT virus";  
RL J. Virol. 75:11811-11820(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;  
RA Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF332093; AAL33392.1; -  
SQ SEQUENCE 321 AA; 35834 MW; 9D9083F6A1BE3E08 CRC64;  
SO  
Query Match 1.8%; Score 8; DB 12; Length 321;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Cy 436 GDDDDDD 443  
Db 170 GDDDDDD 177  
RESULT 72  
091L82 PRELIMINARY; PRT; 321 AA.  
AC 091L82;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE ORF170 (WSSV449).  
OS White spot syndrome virus (WSSV).  
OC Viruses; unclassified viruses.  
OX NCBI\_Taxid=92652;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21342572; PubMed=11448154;  
RA van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,  
RA Tarchini R., Fiers M., Sandbink H., Lankhorst R.K., Vlak J.M.;  
RT "The white spot syndrome virus DNA genome sequence";  
RL Virology 286:7-22(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,  
RA Tarchini R., Fiers M., Sandbink H., Lankhorst R.K., Vlak J.M.;  
RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20517548; PubMed=11062040;  
RA Tsai M.F., Yu H.T., Tzeng H.F., Liu J.H., Chou C.M., Huang C.J.,  
Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;  
RT "Identification and characterization of a shrimp white spot syndrome  
RT virus (WSSV) gene that encodes a novel chimeric polypeptide of  
RT cellular-type thymidine kinase and thymidylate kinase";  
RL Virology 277:100-110(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21844071; PubMed=11853398;  
RA Chen L.L., Liu J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,  
Lo C.F., Kou G.H.;  
RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white  
RT spot syndrome virus and characterization of the motif important for  
RT targeting VP35 to the nuclei of transfected insect cells";  
RL Virology 293:44-53(2002).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX STRAIN-TAIWAN;

RA Lo C.-F., Kou G.-H.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF49029; AAK77839.1; -  
DR EMBL: AF440570; AAL89317.1; -  
SQ SEQUENCE 321 AA; 35833 MW; C39083F6A2DEE805 CRC64;  
SO  
Query Match 1.8%; Score 8; DB 12; Length 321;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Cy 436 GDDDDDD 443  
Db 170 GDDDDDD 177  
RESULT 73  
08S667 PRELIMINARY; PRT; 325 AA.  
AC 08S667;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Putative protein kinase.  
GN OSJNB40034B05.26.  
OS Oryza sativa (japonica cultivar-group). Embryophyta; Tracheophyta;  
OC Eukaryota; Viridiplantae; Streptophyta; Liliopsida; Poales; Poaceae;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Eriactoidae; Oryzae; Oryza.  
OX NCBI\_Taxid=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,  
RA Saski C., Henry D., Oates R., Simmons J.;  
RT "Rice Genomic Sequence";  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC097446; AAM14697.1; -  
KM Kinase.  
SQ SEQUENCE 325 AA; 35532 MW; 415B37F437435D0 CRC64;  
SO  
Query Match 1.8%; Score 8; DB 10; Length 325;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Cy 147 LKIDFGA 154  
Db 182 LKIDFGA 189  
RESULT 74  
08S672 PRELIMINARY; PRT; 326 AA.  
AC 08S672;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Putative protein kinase.  
GN OSJNB40034B05.21.  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Eriactoidae; Oryzae; Oryza.  
OX NCBI\_Taxid=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,  
RA Saski C., Henry D., Oates R., Simmons J.;  
RT "Rice Genomic Sequence";  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC097446; AAM14692.1; -  
KM Kinase.  
SQ SEQUENCE 326 AA; 35591 MW; 98F6F2BAGB0FEF36 CRC64;  
SO  
Query Match 1.8%; Score 8; DB 10; Length 326;

Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 147 LKICDFGA 154  
|||  
Db 183 LKICDFGA 190

## RESULT 75

OBTIF9

ID OBTIF9 PRELIMINARY; PRT; 335 AA.

AC OBTIF9;

DT 01-JUN-2002 (TREMblrel. 21, Created)

DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)

DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)

OS Hypothetical 38.8 KDa protein.

OC Dictyostelium discoideum (Slime mold).

CX Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

OX NCBI\_TaxID=44689;

[1]

SEQUENCE FROM N.A.

STRAIN=AX4;

RA Gloeckner G., Eichinger L., Szafrański K., Pachebat J., Dear P.,

RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,

RA Tungsai B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.,

RT "Sequence and Analysis of Chromosome 2 of Dictyostelium.";

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL/ AC116958; AM08439.1; -.

KM Hypothetical protein.

SQ SEQUENCE 335 AA; 38820 MW; 1B274330A885CF3 CRC64;

Query Match 1.8%; Score 8; DB 5; Length 335;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG:444  
|||  
Db 51 DDDDDDDG 58

Search completed: May 1, 2003, 20:55:30  
Job time : 116 secs



GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:38:47 ; Search time 40 Seconds

(without alignments)  
1515.726 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 455

Sequence: 1 MSSIGAFVQIKFDDLOFFE.....GDDDDDDGEEEDNDWNSB 455

Scoring table:

Gap60 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

rd size: 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database:

A\_Geneseq\_101002.\*  
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	455	100.0	455	21	AA18657 A human regulator
2	455	100.0	455	21	AA183278 Human survival reg
3	455	100.0	455	21	AA183278 Human survival reg
4	447	98.2	473	22	AA183278 Human survival reg
5	331	72.7	800	22	AA183278 Human survival reg
6	331	72.7	800	22	AA183278 Human survival reg
7	269	59.1	349	22	AA183278 Human survival reg
8	124	27.3	124	22	AA183278 Human survival reg
9	118	25.9	141	21	AA183278 Human survival reg
10	80	17.6	92	22	AA183278 Human survival reg

11	80	17.6	92	22	AA17234 Novel human diene
12	56	12.3	144	22	ABG06092 Novel human diene
13	14	3.1	14	21	AA183280 Human survival reg
14	13	2.9	13	22	AA183280 Human survival reg
15	11	2.4	977	22	AA183280 Human survival reg
16	10	2.2	598	11	AA183280 Human survival reg
17	10	2.2	729	11	AA183280 Human survival reg
18	10	2.2	764	17	AA183280 Human survival reg
19	10	2.2	764	17	AA183280 Human survival reg
20	10	2.2	1089	17	AA183280 Human survival reg
21	10	2.2	1089	20	AA183280 Human survival reg
22	10	2.2	1089	21	AA183280 Human survival reg
23	10	2.2	1089	21	AA183280 Human survival reg
24	9	2.0	133	21	AA183280 Human survival reg
25	9	2.0	156	21	AA183280 Human survival reg
26	9	2.0	201	21	AA183280 Human survival reg
27	9	2.0	224	21	AA183280 Human survival reg
28	9	2.0	502	15	AA183280 Human survival reg
29	9	2.0	502	15	AA183280 Human survival reg
30	9	2.0	505	14	AA183280 Human survival reg
31	9	2.0	505	17	AA183280 Human survival reg
32	9	2.0	608	23	AA183280 Human survival reg
33	9	2.0	652	23	AA183280 Human survival reg
34	9	2.0	652	23	AA183280 Human survival reg
35	9	2.0	656	23	AA183280 Human survival reg
36	9	2.0	909	22	AA183280 Human survival reg
37	9	2.0	2391	15	AA183280 Human survival reg
38	8	1.8	20	22	AA183280 Human survival reg
39	8	1.8	20	22	AA183280 Human survival reg
40	8	1.8	21	22	AA183280 Human survival reg
41	8	1.8	21	22	AA183280 Human survival reg
42	8	1.8	21	22	AA183280 Human survival reg
43	8	1.8	22	22	AA183280 Human survival reg
44	8	1.8	22	22	AA183280 Human survival reg
45	8	1.8	22	22	AA183280 Human survival reg
46	8	1.8	22	22	AA183280 Human survival reg
47	8	1.8	22	22	AA183280 Human survival reg
48	8	1.8	22	22	AA183280 Human survival reg
49	8	1.8	22	22	AA183280 Human survival reg
50	8	1.8	22	22	AA183280 Human survival reg
51	8	1.8	22	22	AA183280 Human survival reg
52	8	1.8	22	22	AA183280 Human survival reg
53	8	1.8	22	22	AA183280 Human survival reg
54	8	1.8	22	22	AA183280 Human survival reg
55	8	1.8	22	22	AA183280 Human survival reg
56	8	1.8	22	22	AA183280 Human survival reg
57	8	1.8	22	22	AA183280 Human survival reg
58	8	1.8	22	22	AA183280 Human survival reg
59	8	1.8	22	22	AA183280 Human survival reg
60	8	1.8	22	22	AA183280 Human survival reg
61	8	1.8	22	22	AA183280 Human survival reg
62	8	1.8	22	22	AA183280 Human survival reg
63	8	1.8	22	22	AA183280 Human survival reg
64	8	1.8	22	22	AA183280 Human survival reg
65	8	1.8	22	22	AA183280 Human survival reg
66	8	1.8	22	22	AA183280 Human survival reg
67	8	1.8	22	22	AA183280 Human survival reg
68	8	1.8	22	22	AA183280 Human survival reg
69	8	1.8	22	22	AA183280 Human survival reg
70	8	1.8	22	22	AA183280 Human survival reg
71	8	1.8	22	22	AA183280 Human survival reg
72	8	1.8	22	22	AA183280 Human survival reg
73	8	1.8	22	22	AA183280 Human survival reg
74	8	1.8	22	22	AA183280 Human survival reg
75	8	1.8	22	22	AA183280 Human survival reg
76	8	1.8	22	22	AA183280 Human survival reg
77	8	1.8	22	22	AA183280 Human survival reg
78	8	1.8	22	22	AA183280 Human survival reg
79	8	1.8	22	22	AA183280 Human survival reg
80	8	1.8	22	22	AA183280 Human survival reg
81	8	1.8	22	22	AA183280 Human survival reg
82	8	1.8	22	22	AA183280 Human survival reg
83	8	1.8	22	22	AA183280 Human survival reg



QY 61 SHRNIIQFYGVILPBNVGIYTEYASLGSLYDIYINSNRSEEMDMHIMTAVAKGMHY 120  
 DB 61 SHRNIIQFYGVILPBNVGIYTEYASLGSLYDIYINSNRSEEMDMHIMTAVAKGMHY 120  
 QY 121 LHMARPVKVHRLDKSRNVVIAADGVKICDPGASRPHNHTTHMSLVGTPPMVAPEVIQS 180  
 DB 121 LHMARPVKVHRLDKSRNVVIAADGVKICDPGASRPHNHTTHMSLVGTPPMVAPEVIQS 180  
 QY 181 LPVSETCDTYSYGVVLEWMLTREVPFPGKLEGLQVAMLVKERNLRTIPSSCPSPFAELH 240  
 DB 181 LPVSETCDTYSYGVVLEWMLTREVPFPGKLEGLQVAMLVKERNLRTIPSSCPSPFAELH 240  
 QY 241 QCEWADAKKRPSPFKQIISILESMSNDTSLPDKCNFLHNKAEWRCETIATLERLKLKLERD 300  
 DB 241 QCEWADAKKRPSPFKQIISILESMSNDTSLPDKCNFLHNKAEWRCETIATLERLKLKLERD 300  
 QY 301 LSPFEQELKERERRLKWEOKLTEQSNTPLLPLPLAARMSEESYFESKTESNSAEMSCQI 360  
 DB 301 LSPFEQELKERERRLKWEOKLTEQSNTPLLPLPLAARMSEESYFESKTESNSAEMSCQI 360  
 QY 361 TATSNGEGHGMNPSLOAMLMGFGDIFSMNKAGAVHSGQIMQAKONSKTTSKRGRK 420  
 DB 361 TATSNGEGHGMNPSLOAMLMGFGDIFSMNKAGAVHSGQIMQAKONSKTTSKRGRK 420  
 QY 421 KVNMAIGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455  
 DB 421 KVNMAIGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455

RESULT 2  
 AAY83278  
 ID AAY83278 standard; Protein; 455 AA.  
 AC AAY83278;  
 XX  
 DT 16-AUG-2000 (first entry)  
 DE Human survival regulating kinase (SRK).  
 XX  
 KW Survival regulating kinase; SRK; Raf; MAPK; BAD; MAPKK; MAPKKK; MAPKKL; protein kinase; autophosphorylation; cell growth; regulation; apoptosis; cell survival; nuclear targeting; tumour; human; autoimmune disease.  
 OS Homo sapiens.  
 PN WO200022142-A2.  
 XX  
 DB 20-APR-2000.  
 PF 20-SEP-1999; 99WO-US22008.  
 XX  
 BR 13-OCT-1998; 98US-0104088.  
 XX  
 PA (ONYX-) ONYX PHARM INC.  
 XX  
 PI Ruggieri R, Callow M, Diaz P;  
 XX  
 DR MPI; 2000-317994/27.  
 XX  
 N-PSDB; AAZ93783.  
 XX  
 PT Novel human survival regulating kinase polypeptide for screening agents  
 XX  
 PS Claim 4; Figure 2; 62pp; English.  
 CC Survival regulating kinases (SRK) are a class of proteins involved in  
 CC cell signal transduction pathways such as mitogen-activated protein  
 CC kinase pathways. A protein kinase activity means that the SRK can  
 CC catalyze a reaction in which a phosphate group is transferred from a  
 CC phosphate donor to a phosphate acceptor amino acid residue,  
 CC preferably the hydroxyl side chain of a serine or threonine.

CC Substrates for SRK include SRK, MAPK and BAD and SRK's protein kinase  
 CC activity is similar to that of a MAPKK such as Raf. has a range of  
 CC other activities including a cell growth-regulatory activity, a cell  
 CC survival promoting activity, a HAX-1 binding activity, an apoptosis  
 CC suppressing activity, a MAPKK activation or stimulatory activity, a  
 CC nuclear targeting activity and a SRK-specific immunogenic activity.  
 CC SRK is useful for identifying agents which modulate cellular  
 CC transformations mediated by Ras and SRK and agents that modulate the  
 CC apoptosis suppression activity of SRK. This information may be useful  
 CC in the treatment of autoimmune diseases, tumours and apoptosis  
 CC related disorders.

Sequence 455 AA;  
 SQ  
 Query Match 100.0%; Score 455; DB 21; Length 455;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFVQIKFPDLOFFENGCGGSGSVYRAKVIQDKEVAVKLLKEKEAIIISVL 60  
 DB 1 MSSLGASFVQIKFPDLOFFENGCGGSGSVYRAKVIQDKEVAVKLLKEKEAIIISVL 60  
 QY 61 SHRNIIQFYGVILPBNVGIYTEYASLGSLYDIYINSNRSEEMDMHIMTAVAKGMHY 120  
 DB 61 SHRNIIQFYGVILPBNVGIYTEYASLGSLYDIYINSNRSEEMDMHIMTAVAKGMHY 120  
 QY 121 LHMARPVKVHRLDKSRNVVIAADGVKICDPGASRPHNHTTHMSLVGTPPMVAPEVIQS 180  
 DB 121 LHMARPVKVHRLDKSRNVVIAADGVKICDPGASRPHNHTTHMSLVGTPPMVAPEVIQS 180  
 QY 181 LPVSETCDTYSYGVVLEWMLTREVPFPGKLEGLQVAMLVKERNLRTIPSSCPSPFAELH 240  
 DB 181 LPVSETCDTYSYGVVLEWMLTREVPFPGKLEGLQVAMLVKERNLRTIPSSCPSPFAELH 240  
 QY 241 QCEWADAKKRPSPFKQIISILESMSNDTSLPDKCNFLHNKAEWRCETIATLERLKLKLERD 300  
 DB 241 QCEWADAKKRPSPFKQIISILESMSNDTSLPDKCNFLHNKAEWRCETIATLERLKLKLERD 300  
 QY 301 LSPFEQELKERERRLKWEOKLTEQSNTPLLPLPLAARMSEESYFESKTESNSAEMSCQI 360  
 DB 301 LSPFEQELKERERRLKWEOKLTEQSNTPLLPLPLAARMSEESYFESKTESNSAEMSCQI 360  
 QY 361 TATSNGEGHGMNPSLOAMLMGFGDIFSMNKAGAVHSGQIMQAKONSKTTSKRGRK 420  
 DB 361 TATSNGEGHGMNPSLOAMLMGFGDIFSMNKAGAVHSGQIMQAKONSKTTSKRGRK 420  
 QY 421 KVNMAIGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455  
 DB 421 KVNMAIGFSDFDLSEGDGDDDDGDEEDNDMDNSE 455

RESULT 3  
 AAY84321  
 ID AAY84321 standard; Protein; 455 AA.  
 AC AAY84321;  
 XX  
 DT 12-JUL-2000 (first entry)  
 DE A human cardiovascular system associated protein kinase-2.  
 XX  
 KW Human; cardiovascular system associated protein kinase-2; GAPK-2;  
 KW signalling pathway; cell growth; cell differentiation; gene mapping;  
 KW tissue typing; forensic identification; cardiovascular disease;  
 KW congestive heart failure; transgenic animal.  
 OS Homo sapiens.  
 PN WO200014212-A1.  
 XX  
 DB 16-MAR-2000.  
 PF 09-SEP-1999; 99WO-US20631.

AC	AAW25322;
DT	16-OCT-2001 (first entry)
XX	
DE	Human protein sequence SEQ ID NO:837.
KW	Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW	antihistaminic; antirheumatic; antiarthritic; immunosuppressive;
KW	antibacterial; endocrine; cardiant; central nervous system; virocid;
KW	anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW	antigregant; haemostatic; vulnery; antiviral; osteoporotic; eczema;
KW	dermatologic; antiallergic; antisthmatic; antiparkinsonian; infection;
KW	neuroprotective; antidepressant; nootropic; antidiabetic; inflammation;
KW	immunostimulant; gene therapy; antisense therapy; vaccine; pancreatitis;
KW	antiadysfunction; rheumatoid arthritis; septic shock; pancreatic;
KW	cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW	genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW	thrombocytopenia; osteoporosis; severe combined immunodeficiency;
KW	allergic rhinitis; diabetes; multiple sclerosis; depression;
KW	Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
OS	neurological disorder.
XX	
XX	Homo sapiens.
PX	
PN	WO200153455-A2.
PD	
PF	26-JUL-2001.
PR	22-DEC-2000; 2000WO-US35017.
PR	23-DEC-1999; 99US-0471275.
PR	21-JAN-2000; 2000US-0488725.
PR	25-APR-2000; 2000US-0552317.
PA	(HYSE-) HYSEQ INC.
PI	Tang YT, Liu C, Drmanac RT;
DR	WPI; 2001-457603/49.
DR	N-PDB; AAH92923.
PT	Isolated human polynucleotides encoding polypeptides, useful for the
PT	treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
PS	Claim 20; Page 191; 1217P; English.
XX	
CC	AAH99166 to AAH99904 encode the human proteins given in AAW25225 to
CC	AAH99963. The proteins can have activities based on the tissues and
CC	cells they are expressed in, such as: antineoplastic; antirheumatic;
CC	antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC	central nervous system; virocid; anti-HIV; fungicide; antimutagen;
CC	cardiovascular; antianaemic; antiaggregant; haemostatic; antisthmatic;
CC	antidiabetic; osteopathic; dermatological; antiallergic; antidiabetic;
CC	antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC	encoding them can be used in gene therapy, antisense therapy and vaccine
CC	production. The proteins and polynucleotides are useful for screening for
CC	agonists or antagonists of a protein and for the treatment and diagnosis
CC	of disorders associated with the activity of a protein e.g. inflammation,
CC	rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC	neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC	infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC	anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC	osteoporosis, severe combined immunodeficiency, eczema, allergic
CC	rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC	Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC	neurological disorders.
SQ	Sequence 473 AA;
Query Match	98.2%; Score 447; DB 22; Length 473;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 447; Conservative	0; Mismatches 0; Indels 0; Gaps 0;



QY 1 MSSLGASFWOIKFDDLOFFENCSSGSGSVYRAKWIISODKEVAVKLLKIEKAEIILSVL 60  
 DB 19 MSSLGASFWOIKFDDLOFFENCSSGSGSVYRAKWIISODKEVAVKLLKIEKAEIILSVL 78  
 QY 61 SHRNIIQFYGVILLEPPNYGIVTEVASLGSIVDYINSNSSEMDMDHMTATDVAKGMHY 120  
 DB 79 SHRNIIQFYGVILLEPPNYGIVTEVASLGSIVDYINSNSSEMDMDHMTATDVAKGMHY 138  
 QY 121 LHMEAPVKVIRHDLKSRNVVIAADGVLCICDFGASRFNNHTTHMSLVGTFFPMMAPEVIQS 180  
 DB 139 LHMEAPVKVIRHDLKSRNVVIAADGVLCICDFGASRFNNHTTHMSLVGTFFPMMAPEVIQS 198  
 QY 181 LPVSETCTDYISYGVVLMEMLTREVPEFKGLEGLQVAMLVENKNERLTIIPSSCPSRPAELIH 240  
 DB 199 LPVSETCTDYISYGVVLMEMLTREVPEFKGLEGLQVAMLVENKNERLTIIPSSCPSRPAELIH 258  
 QY 241 QCWEADAKKRPSFKQIISILSMSNDTSLPDKCNSFLHNKAEMRCEIATLERLKKLERD 300  
 DB 259 QCWEADAKKRPSFKQIISILSMSNDTSLPDKCNSFLHNKAEMRCEIATLERLKKLERD 318  
 QY 301 LSFKEQELKERERRLKMEQKLTQOSNTPLLPLAAMSEBSYFESKTEESNAEMSCOI 360  
 DB 319 LSFKEQELKERERRLKMEQKLTQOSNTPLLPLAAMSEBSYFESKTEESNAEMSCOI 378  
 QY 361 TATNGEGHGNPSTLOAMMLMGFGDIFSNKKAQVNHSGQIINQAKONSSKTTSKRRGK 420  
 DB 379 TATNGEGHGNPSTLOAMMLMGFGDIFSNKKAQVNHSGQIINQAKONSSKTTSKRRGK 438  
 QY 421 KVNNALGPSDFDLSGGDDDDDDDEEE 447  
 DB 439 KVNNALGPSDFDLSGGDDDDDDDEEE 465

## RESULT 5

ID AAB71957 standard; Protein; 800 AA.  
 AC AAB71957;  
 DT 11-MAY-2001 (first entry)  
 DE Human TGF-beta receptor encoded by cDNA clone HDP5W48.  
 KW Human; antiinflammatory; anti-HIV; immunostimulant; cardiac; vascular; ophthalmological; neuroprotective; nociceptive; anticonvulsant; vaccine; antiparkinsonian; antimicrobial; vulnery; gene therapy; infection; transforming growth factor; TGF; TGF-beta receptor; immune disorder; hyperproliferative disorder; cardiovascular disease; angiogenesis;  
 KW neurological disorder.  
 OS Homo sapiens.  
 PN MO200112670-A1.  
 PD 22-FEB-2001.  
 PF 10-AUG-2000; 2000WO-US21736.  
 PR 13-AUG-1999; 99US-0148682.  
 PR 20-SEP-1999; 99US-0154887.  
 PA (HUMA-) HUMAN GENOME SGT INC.  
 PI Ruben SM, NI J;  
 WI: 2001-202858/20.  
 N-PSDB; AAF75336.  
 PT Nucleic acid molecules encoding 12 transforming growth factor-beta receptor polypeptides, useful for preventing, diagnosing and treating e.g. cancers, Parkinson's disease and diabetic retinopathy -

XX Claim 11; Page 293-295; 311pp; English.

CC The present sequence is one of 12 novel human transforming growth factor (TGF)-beta receptor polypeptides. The TGF-beta receptor polypeptides and polypeptides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate polypeptide expression. Such diseases include immune disorders (e.g. multiple sclerosis, systemic lupus erythematosus and human immunodeficiency virus (HIV) infections), hyperproliferative disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases (e.g. Schmitz syndrome, Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic disorders (e.g. corneal graft neovascularization and diabetic retinopathy), neurological disorders (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease) and infectious diseases. The polypeptides and polypeptides are also useful for promoting wound healing, regeneration and/or chemotaxis. The polynucleotides and their complementary sequences may also be used as probes in diagnostic assays to detect and quantify the presence of similar nucleic acid sequences in samples. The polypeptides may be used as antigens in the production of antibodies and in assays to identify modulators of protein expression and activity. The anti-TGF-beta receptor antibodies may be used to down regulate expression and activity and as diagnostic agents for detecting the presence of the polypeptides in samples.

CC Sequence 800 AA;

Query Match 72.7%; Score 331; DB 22; Length 800;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-304;

Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLGASFWOIKFDDLOFFENCSSGSGSVYRAKWIISODKEVAVKLLKIEKAEIILSVL 60  
 DB 1 MSSLGASFWOIKFDDLOFFENCSSGSGSVYRAKWIISODKEVAVKLLKIEKAEIILSVL 60  
 QY 61 SHRNIIQFYGVILLEPPNYGIVTEVASLGSIVDYINSNSSEMDMDHMTATDVAKGMHY 120  
 DB 61 SHRNIIQFYGVILLEPPNYGIVTEVASLGSIVDYINSNSSEMDMDHMTATDVAKGMHY 120  
 QY 121 LHMEAPVKVIRHDLKSRNVVIAADGVLCICDFGASRFNNHTTHMSLVGTFFPMMAPEVIQS 180  
 DB 121 LHMEAPVKVIRHDLKSRNVVIAADGVLCICDFGASRFNNHTTHMSLVGTFFPMMAPEVIQS 180  
 QY 181 LPVSETCTDYISYGVVLMEMLTREVPEFKGLEGLQVAMLVENKNERLTIIPSSCPSRPAELIH 240  
 DB 181 LPVSETCTDYISYGVVLMEMLTREVPEFKGLEGLQVAMLVENKNERLTIIPSSCPSRPAELIH 240  
 QY 241 QCWEADAKKRPSFKQIISILSMSNDTSLPDKCNSFLHNKAEMRCEIATLERLKKLERD 300  
 DB 241 QCWEADAKKRPSFKQIISILSMSNDTSLPDKCNSFLHNKAEMRCEIATLERLKKLERD 300  
 QY 301 LSFKEQELKERERRLKMEQKLTQOSNTPLL 331  
 DB 301 LSFKEQELKERERRLKMEQKLTQOSNTPLL 331

## RESULT 6

ID AAB65673 standard; Protein; 800 AA.  
 AC AAB65673;  
 DT 27-MAR-2001 (first entry)  
 DE Novel protein kinase, SEQ ID NO: 201.  
 KW Human; mouse; protein kinase; antiarthritic; antiinflammatory; osteopathic; dermatological; antiinflammatory; renal; antiinflammatory; antiasthmatic; immunosuppressive; cardiac; renal; antiinflammatory; antiasthmatic; immune disorder; cardiovascular disease; neurodegenerative disease; cancer; autoimmune disorder; stroke; inflammatory bowel disease; inflammatory pelvic disease; multiple sclerosis; psoriasis.



Qy 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLSLQVAMLVENKNERLTTPSSCCPSFAELH 240  
 Db 237 LPVSETCDTYSYGVVLMEMLTREVPFKGLSLQVAMLVENKNERLTTPSSCCPSFAELH 236  
 Qy 241 QCWEADAKKRPSPFOIITILESMSNDLSL 269  
 Db 297 QCWEADAKKRPSPFOIITILESMSNDLSL 325

RESULT 8  
 AAB65552  
 ID AAB65552 standard; Protein; 124 AA.  
 AC AAB65552;  
 XX  
 XX 27-MAR-2001 (first entry)  
 DE C-terminus specific to novel human protein kinase MLK4B.

Human; protein kinase; antirheumatic; antisclerotic; immunosuppressive;  
 dermant; renal; antiinflammatory; antidiabetic; osteopathic;  
 dermatological; antidiabetic; antiinfectivity; gene therapy; vaccine;  
 immune disorder; cardiovascular disease; neurodegenerative disease;  
 cancer; autoimmune disorder; stroke; inflammatory bowel disease;  
 inflammatory pelvic disease; multiple sclerosis; psoriasis.

OS Homo sapiens.  
 XX WO200073469-A2.  
 PN 07-DEC-2000.  
 PD 26-MAY-2000; 2000WO-US14842.  
 PF 28-MAY-1999; 99US-0136503.  
 PR (SUGEN-) SUGEN INC.  
 PA Plowman GD, Martinez R, Whyte D, Sudersanam S;  
 PI WPI; 2001-032161/04.

DR Nucleic acids encoding kinase polypeptides, useful for diagnosing and  
 XX treating immune-related diseases and disorders, cardiovascular disease,  
 PT neurodegenerative diseases and/or cancers -  
 PS Example 1; Page 116; 310pp; English.

The present sequence is given in a specification relating to novel  
 protein kinases. The protein kinases and the nucleic acids that encode  
 them may be used in the treatment and diagnosis of diseases  
 associated with inappropriate kinase expression such as immune-related  
 diseases and disorders, cardiovascular disease, neurodegenerative  
 diseases and/or cancers. The nucleic acids and complementary sequences  
 may also be used as DNA probes in diagnostic assays. The kinase  
 polypeptides may be used as antigens in the production of antibodies of  
 kinase expression and activity. Anti-kinase antibodies and kinases  
 antagonists may also be used to down regulate kinase expression and  
 activity. Diseases related to kinase expression and activity include  
 rheumatoid arthritis, atherosclerosis, autoimmune disorders,  
 complications of organ transplantation, myocardial infarction, immune  
 disorders, cardiomypathies, strokes, renal failure, oxidative-stress  
 related disorders, chronic inflammatory bowel disease, chronic  
 inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis,  
 psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive  
 disorders.

SO Sequence 124 AA;  
 Query Match 27.3%; Score 124; DB 22; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-109;  
 Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 332 LPLAARMESESYFESKTEESNAEMSCOTITATSGEGHGNPSIQAMLMGFGDIFSMNK 391  
 Db 1 LPLAARMESESYFESKTEESNAEMSCOTITATSGEGHGNPSIQAMLMGFGDIFSMNK 60  
 Qy 392 AGAVHSGMQINMOAKONSSKTSKRGGKVMALGSPDPLSGDDDDDDDDGDEENDM 451  
 Db 61 AGAVHSGMQINMOAKONSSKTSKRGGKVMALGSPDPLSGDDDDDDDDGDEENDM 120  
 Qy 452 DNSB 455  
 Db 121 DNSB 124

RESULT 9  
 AAG03583  
 ID AAG03583 standard; Protein; 141 AA.  
 AC AAG03583;  
 XX 06-OCT-2000 (first entry)  
 DE Human secreted protein, SEQ ID NO: 7664.

Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 gene therapy; chromosome mapping.  
 OS Homo sapiens.  
 XX EP1033401-A2.  
 PN 06-SEP-2000.  
 PD 21-FEB-2000; 2000EP-0200610.  
 PF 26-FEB-1999; 99US-0122487.  
 PR (GENST) GENSET.

PA Dumas Milne Edwards J, Duclert A, Giordano J;  
 PI WPI; 2000-500381/45.  
 DR N-PSDB; AAC03589.  
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 PS Claim 13; SEQ ID 7664; 71pp + CD-ROM; English.

The present sequence is a polypeptide encoded by one of a large number  
 of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
 were prepared from total human RNAs or polyA+ RNAs derived from 30  
 different tissues. EST sequences usually correspond mainly to the 3'  
 untranslated region (UTR) of the mRNA because they are often obtained  
 from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
 isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 those cases where longer cDNA sequences have been obtained, the full 5'  
 UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 ends and can therefore be used to obtain full length cDNAs and genomic  
 DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 chromosome mapping procedures. They are used to obtain upstream  
 regulatory sequences and to design expression and secretion vectors.

SO Sequence 141 AA;  
 Query Match 25.9%; Score 118; DB 21; Length 141;  
 Best Local Similarity 100.0%; Pred. No. 2e-103;  
 Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 102 MDMDHIMTWADVAKGMHYLHMEAPVKYIHRDLKSRNVVIAADGVLCICPGASRPHNT 161  
 Db 1 MDMDHIMTWADVAKGMHYLHMEAPVKYIHRDLKSRNVVIAADGVLCICPGASRPHNT 60

QY 162 THMSLNGTFPMMAPEVIOQLPVSETCDTYSXGVVLMETREVEFKGLEIQAVALIV 219  
DB 61 THMSLVGTFPMMAPEVIOQLPVSETCDTYSXGVVLMETREVEFKGLEIQAVALIV 118  
RESULT 10  
AAU87295 \*  
ID AAU87295 standard; Protein: 92 AA.  
XX  
AC AAU87295;  
XX  
DT 05-JUN-2002 (first entry)  
XX  
DE Novel central nervous system protein #205.  
XX  
KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;  
KW hyperproliferative disorder; neoplasia; cardiovascular disorder;  
KW cardiac arrest; cerebrovascular disorder; ischemia; angiogenesis;  
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;  
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;  
KW adenocarcinoma; reproductive system disorder; testicular feminisation;  
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;  
KW respiratory disorder; renal disorder; kidney failure; blood disorder;  
KW myocardial infarction; wound healing; cell proliferation; skin aging;  
KW food additive; food preservative; gene therapy.  
XX  
OS Homo sapiens.  
XX  
EN WO20015318-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01332.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214686.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226271.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0233397.  
PR 14-SEP-2000; 2000US-0233398.  
PR 14-SEP-2000; 2000US-0233399.  
PR 14-SEP-2000; 2000US-0233400.  
PR 14-SEP-2000; 2000US-0233401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 25-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239335.  
PR 13-OCT-2000; 2000US-0239337.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.



```

PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 01-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.

```

```

PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0251989.
PR 06-DEC-2000; 2000US-0251990.
PR 06-DEC-2000; 2000US-0251997.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM,
XX WPI, 2001-465460/50.
XX N-PSDB; AAS27151.
XX
XX Novel polypeptides useful for diagnosing, treating, preventing and/or
XX PT prognosing disorders related to the proteins, including cancers, immune
XX PT disorders and neuronal disorders
XX
XX PS Claim 1; SEQ ID No 799; 880bp; English.
XX
XX The invention relates to novel isolated polypeptides (I), and
XX CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
XX CC diagnosing, preventing and treating diseases including immune system
XX CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
XX CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
XX CC transplant rejections and graft versus host disease, infectious diseases
XX CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
XX CC other blood-related disorders (sickle cell anaemia), myeloproliferative
XX CC disorders, primary haematopoietic disorders, hyperproliferative
XX CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
XX CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
XX CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
XX CC disorders (e.g. glomerulonephritis), cardiovascular disorders, in
XX CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
XX CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
XX CC Addison's disease), reproductive system disorders, gastrointestinal
XX CC disorder (inflammatory disorders), liver disorders (cirrhosis),
XX CC as stimulators of B-cell responsiveness to pathogens, activators of
XX CC T-cells, to induce higher affinity antibodies, and as a means to induce
XX CC tumour proliferation in pathologies e.g. acquired immune deficiency
XX CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction
XX CC pathway protein, amino acid sequences of the invention.
XX
XX
XX Query Match 17.6%; Score 80; DB 22; Length 92;
XX Best Local Similarity 100.0%; Pred. No. 1.3e-67;
XX Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 54 AELLVLSHRNIIQFYGVILEPPNYGIVTEYASIGSLDYINRSRSEBMDHMTATD 113
XX |||||
XX Db 13 AELLVLSHRNIIQFYGVILEPPNYGIVTEYASIGSLDYINRSRSEBMDHMTATD 72
XX |||||
XX QY 114 VAKGMHYLMEAPVKVIRHD 133
XX |||||
XX Db 73 VAKGMHYLMEAPVKVIRHD 92
XX |||||
XX
XX RESULT 12

```

ABG06092  
ID ABG06092 standard; Protein, 144 AA.  
AC ABG06092;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #6083.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US08631.  
XX  
31-MAR-2000; 2000US-0540217.  
23-AUG-2000; 2000US-0649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI; 2001-639362/73.  
XX  
DR N-PSDB; AAS70279.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders of other traits and to assess  
PT biodiversity  
XX  
PS Claim 20; SEQ ID No 36451; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridization probes,  
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX and gene mapping, and in recombinant production of (II). The  
XX polynucleotides are also used in diagnostics as expressed sequence tags  
XX for identifying expressed genes. (I) is useful in gene therapy techniques  
XX (II). (II) is useful for generating antibodies against it, detecting or  
XX quantitating a polypeptide in tissue, as molecular weight markers and as  
XX a food supplement. (II) and its binding partners are useful in medical  
XX imaging of sites expressing (II). (I) and (II) are useful for treating  
XX disorders involving aberrant protein expression or biological activity.  
XX The polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. ABG00010-ABG03037 represent novel human  
XX diagnostic amino acid sequences of the invention.  
XX Note: The sequence data for this patent did not appear in the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.  
SQ Sequence 144 AA;  
Query Match 12.3%; Score 56; DB 22; Length 144;  
Best Local Similarity 100.0%; Pred. No. 9.7e-45;  
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 139 VVIADGVLTICDFGASRPHNHTTMSLVGTFPMWAPVIOSLPVSETCDTYSYGV 194  
ID VVIADGVLTICDFGASRPHNHTTMSLVGTFPMWAPVIOSLPVSETCDTYSYGV 113  
RESULT 13  
AAV83280  
ID AAV83280 standard; Protein, 14 AA.

XX  
AC AAV83280;  
XX  
DT 16-AUG-2000 (first entry)  
XX  
DE Human survival regulating kinase (SRK) antigenic peptide.  
XX  
KW Survival regulating kinase; SRK; Raf; MBP; BAD; MAPKK; MAPKK;  
KW HAX-1; protein kinase; autophosphorylation; cell growth; regulation;  
KW apoptosis; cell survival; nuclear targeting; tumour; human;  
KW autoimmune disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200022142-A2.  
XX  
PD 20-APR-2000.  
XX  
PF 20-SEP-1999; 99WO-US22008.  
XX  
PR 13-OCT-1998; 98US-0104088.  
XX  
PA (ONYX-) ONYX PHARM INC.  
XX  
PI Ruggieri R, Callow M, Diaz P;  
XX  
DR WPI; 2000-317994/27.  
XX  
PT Novel human survival regulating kinase polypeptide for screening agents  
PT which modulate biological pathways associated with SRK useful in  
PT treating autoimmune diseases, tumors and apoptosis-related disorders  
XX  
PS Claim 53; Page 6; 62pp; English.  
XX  
XX Survival regulating kinases (SRK) are a class of proteins involved in  
XX cell signal transduction pathways such as mitogen-activated protein  
XX kinase pathways. A protein kinase activity means that the SRK can  
XX catalyse a reaction in which a phosphate group is transferred from a  
XX phosphate donor to a phosphate acceptor amino acid residue.  
XX preferably the hydroxyl side chain of a serine or threonine.  
XX Substrates for SRK include SRK, MBP and BAD and SRK's protein kinase  
XX activity is similar to that of a MAPKK such as Raf. has a range of  
XX other activities including a cell growth-regulatory activity, a cell  
XX survival promoting activity, a HAX-1 binding activity, an apoptosis  
XX suppressing activity a MAPK activation or stimulatory activity, a  
XX nuclear targeting activity and a SRK-specific immunogenic activity.  
XX SRK is useful for identifying agents which modulate cellular  
XX transformations mediated by Ras and SRK and agents that modulate the  
XX apoptosis suppression activity of SRK. This information may be useful  
XX in the treatment of autoimmune diseases, tumours and apoptosis  
XX related disorders.  
SQ Sequence 14 AA;  
Query Match 3.1%; Score 14; DB 21; Length 14;  
Best Local Similarity 100.0%; Pred. No. 7e-06;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 406 AKONSSKTSKRRG 419  
ID AKONSSKTSKRRG 14  
Db 1 AKONSSKTSKRRG 14  
RESULT 14  
AAV71969  
ID AAV71969 standard; Peptide, 13 AA.  
XX  
AC AAV71969;  
XX  
DT 11-MAY-2001 (first entry)  
XX  
DE Human TGF-beta receptor serine/threonine kinase domain #2.  
XX

Human; antisclerotic; dermatological; immunosuppressive; cytostatic;  
 anti-inflamatory; anti-HIV; immunostimulant; caridiac; vascular;  
 ophthalmological; neuroprotective; nootropic; anticonvulsant; vaccine;  
 antiparkinsonian; antimicrobial; vulnery; gene therapy; infection;  
 transforming growth factor; TGF; TGF-beta receptor; immune disorder;  
 hyperproliferative disorder; cardiovascular disease; angiogenesis;  
 neurological disorder.

Homo sapiens.

MO200112670-A1.

22-FEB-2001.

10-AUG-2000; 2000WO-US21736.

13-AUG-1999; 99US-0148682.

20-SEP-1999; 99US-0154887.

(HUMA-) HUMAN GENOME SCI INC.

Ruben SM, Ni J;

WPI; 2001-202858/20.

Nucleic acid molecules encoding 12 transforming growth factor-beta  
 receptor polypeptides, useful for preventing, diagnosing and treating  
 e.g. cancers, Parkinson's disease and diabetic retinopathy -

Disclosure; Page 308; 31pp; English.

The present sequence is provided in a specification relating to nucleic  
 acid molecules encoding 12 novel human transforming growth factor  
 (TGF) beta receptor polypeptides. The TGF-beta receptor polynucleotides  
 and polypeptides may be used in the prevention, diagnosis and treatment  
 of diseases associated with inappropriate polypeptide expression. Such  
 diseases include immune disorders (e.g. multiple sclerosis, systemic  
 lupus erythematosus and human immunodeficiency virus (HIV) infections),  
 hyperproliferative disorders (e.g. cancers and Gaucher's disease),  
 cardiovascular diseases (e.g. Schmitz syndrome, Chaga's cardiomyopathy  
 and coronary arteriosclerosis), angiogenic disorders (e.g. corneal graft  
 neovascularisation and diabetic retinopathy), neurological disorders  
 (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease)  
 and infectious diseases. The polynucleotides and polypeptides are also  
 useful for promoting wound healing, regeneration and/or chemotaxis. The  
 polynucleotides and their complementary sequences may also be used as DNA  
 probes in diagnostic assays to detect and quantitate the presence of  
 similar nucleic acid sequences in samples. The polypeptides may be used  
 as antigens in the production of antibodies and in assays to identify  
 modulators of protein expression and activity. The anti-TGF-beta receptor  
 antibodies may be used to down regulate expression and activity and as  
 diagnostic agents for detecting the presence of the polypeptides in  
 samples.

Sequence 13 AA;

Query Match 2.9%; Score 13; DB 22; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-05;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 129 VHRDLKSRNVV 141  
 |||||  
 Db 1 VHRDLKSRNVV 13

RESULT 15  
 ID ABB71694  
 AC ABB71694 standard; Protein; 977 AA.  
 AC ABB71694;  
 XX 26-MAR-2002 (first entry)  
 DT  
 XX

Drosophila melanogaster polypeptide SEQ ID NO 41874.

Drosophila; developmental biology; cell signalling; insecticide;  
 pharmaceutical.

Drosophila melanogaster.

MO200171042-A2.

27-SEP-2001.

23-MAR-2001; 2001WO-US09231.

23-MAR-2000; 2000US-191637P.

11-JUL-2000; 2000US-0614150.

(PEKE) PE CORP NY.

Venter JC, Adams M, Li PWD, Myers EW;

WPI; 2001-656860/75.

N-PSDB; ABL15797.

New isolated nucleic acid detection reagent for detecting 1000 or more  
 genes from Drosophila and for elucidating cell signalling and cell-cell  
 interactions -

Disclosure; SEQ ID NO 41874; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent  
 capable of detecting 1000 or more genes from Drosophila. The invention is  
 useful in developmental biology and in elucidating cell signalling and  
 cell-cell interactions in higher eukaryotes for the development of  
 insecticides, therapeutics and pharmaceutical drugs. The invention  
 discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 sequences (ABL01840-ABL16175) and the encoded proteins  
 (ABBS7737-ABBS72072).

The sequence data for this patent did not form part of the printed  
 specification, but was obtained in electronic format directly from WIPO  
 at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 977 AA;

Query Match 2.4%; Score 11; DB 22; Length 977;  
 Best Local Similarity 100.0%; Pred. No. 0.19;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 191 SYGVVLMEMLT 201  
 |||||  
 Db 331 SYGVVLMEMLT 341

RESULT 16  
 ID AAR20470  
 AC AAR20470 standard; Protein; 598 AA.  
 AC AAR20470;  
 DT 15-APR-1992 (first entry)  
 DE Soluble ysc alpha.  
 XX Endoprotease; muten; carboxypeptidase; ysc alpha; KEX1; KEX2.  
 XX Saccharomyces cerevisiae.  
 PN BP467839-A.  
 XX 22-JAN-1992.  
 PD 09-JUL-1991; 91EP-0810543.  
 PF 18-JUL-1990; 90GB-0015825.  
 PR



XX (CIBA ) CIBA GEIGY AG.  
 XX  
 PI Helm J, Seebach P, Takabayashi K;  
 DR WPI, 1992-026641/04.  
 DR N-PSDB; AAQ04768.  
 XX  
 PT In vitro processing of fusion protein - using yeast endo:protease  
 PT ysc F and carboxypeptidase ysc alpha, for produ. of biologically  
 PT active mature protein  
 XX  
 PS Example 8; Page 23; 35pp; English.  
 XX  
 CC The sequence (SEQ ID NO: 2) is encoded by an insert from plasmid  
 CC pDP34/GAPL-KEX1 and is a sol. form of ysc alpha. The KEX1  
 CC gene encoding the complete protein (A. Dmochowska et al., Cell 50,  
 CC 573-584 (1987)) was digested with EcoRI and XhoI to isolate a  
 CC fragment encoding a C-terminal truncated sol. protein. The gene  
 CC was expressed in *S. cerevisiae*. The resulting recombinant sol.  
 CC enzyme can be used, together with similarly produced sol. yscF  
 CC for the prep. of mature proteins by in vitro processing of  
 CC suitably tailored fusion proteins (see AAR0471, 72).  
 CC See also AAR0469-72.  
 CC  
 SQ Sequence 598 AA;  
 XX  
 QY 434 SEGDDDDDD 443  
 DB 564 SEGDDDDDD 573  
 XX  
 RESULT 17  
 ID AAR05457 standard; protein; 729 AA.  
 XX  
 AC AAR05457;  
 XX  
 DT 17-OCT-1990 (first entry)  
 XX  
 DE KEX1 amino acid sequence.  
 XX  
 KW KEX1; carboxypeptidase; protease; insulin; ds.  
 XX  
 SS Saccharomyces cerevisiae.  
 XX  
 PN US4929553-A.  
 XX  
 PD 29-MAY-1990.  
 XX  
 PF 29-MAY-1987; 87US-0055555.  
 XX  
 PR 29-MAY-1987; 87US-0055555.  
 XX  
 PA (CANA ) CANADIAN PATENT & DEV LTD.  
 XX  
 PI Busey H, Dmochowska A, Thomas DY, Dignard D;  
 XX  
 DR WPI; 1990-192933/25.  
 DR N-PSDB; AAQ04768.  
 XX  
 PT Specific processing of secreted proteins in transformed yeast -  
 PT using vector contg. yeast KEX1 gene, encoding protease, and DNA  
 PT sequence encoding precursor polypeptide eg. Preproinsulin  
 PS Disclosure; ; p; English.  
 CC Gene product is a protease with carboxypeptidase B activity, useful  
 CC in transforming yeast, producing gene products such as prepro

CC insulin and prepro opio, melanocortin.  
 XX  
 SQ Sequence 729 AA;  
 XX  
 QY 434 SEGDDDDDD 443  
 DB 564 SEGDDDDDD 573  
 XX  
 RESULT 18  
 ID AAW01897 standard; Protein; 764 AA.  
 XX  
 AC AAW01897;  
 XX  
 DT 24-NOV-1996 (first entry)  
 XX  
 DE Nonsense-mediated mRNA decay 2 C-terminal.  
 XX  
 KW Nonsense-mediated mRNA decay; NMD2; Upflp; inhibition.  
 XX  
 OS Saccharomyces cerevisiae.  
 XX  
 PN WO9622301-A1.  
 XX  
 PD 25-JUL-1996.  
 XX  
 PF 27-DEC-1995; 95MO-US16930.  
 XX  
 PR 20-JAN-1995; 95US-0375300.  
 XX  
 PA (UYMA-) UNIV MASSACHUSETTS MEDICAL CENT.  
 XX  
 PI He F, Jacobson AS;  
 XX  
 DR WPI; 1996-354469/35.  
 DR N-PSDB; AATJ1984.  
 XX  
 PT Isolated nonsense-mediated mRNA decay gene and protein - used to  
 PT develop methods for inhibiting the decay pathway for producing  
 PT heterologous or endogenous proteins  
 XX  
 SS Claim 10; Page 54-56; 73pp; English.  
 XX  
 CC The NMD2 gene (AATJ1993) is named after its role in the Nonsense-  
 CC mediated mRNA Decay pathway. The protein, Nmd2p, binds to Upflp.  
 CC A C-terminal fragment of the protein (AATJ1994) also binds Upflp.  
 CC when overexpressed in the host cell, the fragment inhibits the  
 CC function of Upflp, thereby inhibiting the nonsense-mediated  
 CC mRNA decay pathway.  
 XX  
 SQ Sequence 764 AA;  
 XX  
 QY 437 DDDDDDDGEE 446  
 DB 577 DDDDDDDGEE 586  
 XX  
 RESULT 19  
 ID AAY98056 standard; Protein; 764 AA.  
 XX  
 AC AAY98056;  
 XX  
 DT 07-SEP-2000 (first entry)

XX DE Yeast NMD2 carboxy terminal fragment.

XX XX

KW Yeast; nonsense-mediated mRNA decay pathway; NMD2; UPF3; UPF1;

KW upf1p; gene therapy; beta-thalassemia; cancer;

KW polycystic kidney disease; Duchenne muscular dystrophy; Fanconi anaemia;

KW haemophilia; hypercholesterolaemia; neurofibromatosis; Tay-Sachs disease;

KW glycogen storage disease; cystic fibrosis; adenomatous polyposis coli;

KW Cowden disease; Maple syrup urine disease; Wilson disease;

KW Niemann-Pick disease; Turcot syndrome; McArdle disease;

KW ornithine transcarbamylase deficiency; antisense therapy; ribozyme.

OS Saccharomyces cerevisiae.

XX US6071700-A.

PN 06-JUN-2000.

XX 21-OCT-1998; 98US-0177431.

XX 20-JAN-1995; 95US-0375300.

XX 21-OCT-1997; 97US-0955472.

XX (UYMA-) UNIV MASSACHUSETTS.

PA Jacobson AS, He F;

PI WPI; 2000-422078/36.

XX N-PSDB; AAA39451.

DR

XX Testing for compounds able to modulate the nonsense mediated mRNA decay

PT pathway for increasing endogenous protein production by incubating a

PT candidate compound with a cell having a gene with a nonsense mutation

PT and a gene without mutation -

PS Disclosure; Fig 1; 57pp; English.

XX

XX The nonsense-mediated mRNA decay pathway has an important role in the

CC degradation of mRNA transcripts that contain a stop codon. The products

CC of the UPF1 and UPF3 genes (proteins Upf1p and Upf3p, respectively) are

CC essential components of this pathway. Mutations in these genes stabilise

CC mRNA containing premature stop codons. The present sequence is the

CC carboxy terminal of the NMD2 gene from Saccharomyces cerevisiae. When

CC the present sequence is overexpressed, it binds to Upf1p to inhibit its

CC function. Inhibition of the nonsense-mediated mRNA decay pathway is a

CC useful means of treating disorders caused by the presence of nonsense

CC mutations, e.g. breast cancer, polycystic kidney diseases, Niemann-Pick

CC disease, adenomatous polyposis coli, cystic fibrosis, Fanconi's anaemia,

CC haemophilia, hypercholesterolaemia, neurofibromatosis, ornithine

CC transcarbamylase deficiency, retinoblastoma, glycogen storage disease,

CC McArdle disease, colorectal cancer, Tay-Sachs disease, Cowden disease,

CC Wilson disease, polycystic kidney disease, Duchenne muscular dystrophy,

CC adrenocortical carcinoma, Li-Fraumeni syndrome, Maple syrup urine

CC disease, Turcot syndrome or beta-thalassemia. Antisense mRNA molecules

CC complementary to the NMD2 gene (AAA39450) may be used for antisense

CC therapy of the above mentioned disorders since the antisense molecules

CC would inhibit translation of the NMD2 mRNA. Additionally, the NMD2 gene

CC may be manipulated for ribozyme therapy of the above mentioned

CC disorders: ribozymes designed to catalytically cleave nonsense-mediated

CC mRNA decay pathway mRNAs e.g. NMD2 mRNA, can also be used to prevent

CC translation and therefore expression of these mRNAs.

XX

XX Sequence 764 AA;

XX

Query Match 2.2%; Score 10; DB 21; Length 764;

Best Local Similarity 100.0%; Pred. No. 1.4;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGEE 446

DB 577 DDDDDDDGEE 586

RESULT 20

ID AAW01896 standard; Protein; 1089 AA.

XX AAW01896;

XX

AC AAW01896;

XX

DT 24-NOV-1996 (first entry)

XX

DE Nonsense-mediated mRNA decay 2 protein.

XX

XX Nonsense-mediated mRNA decay; NMD2; Upf1p; inhibition.

KW

XX Saccharomyces cerevisiae.

OS

XX WO9622301-A1.

PN 25-JUL-1996.

XX 27-DEC-1995; 95WO-US16930.

XX 20-JAN-1995; 95US-0375300.

XX (UYMA-) UNIV MASSACHUSETTS MEDICAL CENT.

PA He F, Jacobson AS;

PI WPI; 1996-354469/35.

DR N-PSDB; AAT31993.

XX

XX Isolated nonsense-mediated mRNA decay gene and protein - used to

PT develop methods for inhibiting the decay pathway for producing

PT heterologous or endogenous proteins

XX

PS Claim 8; Page 49-52; 73pp; English.

XX

XX The NMD2 gene (AAT31993) is named after its role in the Nonsense-

CC Mediated mRNA Decay pathway. The protein, Nmd2p, binds to Upf1p.

CC A C-terminal fragment of the protein (AAT31994) also binds Upf1p and,

CC when overexpressed in the host cell, the fragment inhibits the

CC function of Upf1p, thereby inhibiting the nonsense-mediated

CC mRNA decay pathway.

XX

XX Sequence 1089 AA;

XX

QY 437 DDDDDDDGEE 446

DB 902 DDDDDDDGEE 911

Query Match 2.2%; Score 10; DB 17; Length 1089;

Best Local Similarity 100.0%; Pred. No. 1.9;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGEE 446

DB 902 DDDDDDDGEE 911

RESULT 21

ID AAY05835 standard; Protein; 1089 AA.

XX AAY05835;

XX

AC AAY05835;

XX

DT 02-AUG-1999 (first entry)

XX

DE Yeast Nmd2p involved in nonsense-mediated mRNA decay pathway.

XX

XX NMD2 gene; Nmd2p protein; nonsense-mediated mRNA decay pathway;

KW yeast; nonsense mutation; therapy; breast cancer; haemophilia;

KW polycystic kidney disease; Niemann-Pick disease; cystic fibrosis;

KW adenomatous polyposis coli; Fanconi's anaemia; neurofibromatosis;

KW hypercholesterolaemia; ornithine transcarbamylase deficiency;

KW retinoblastoma; glycogen storage disease; McArdle disease; cancer;

KW Tay-Sachs disease; Cowden disease; Wilson disease;

KW beta-thalassemia.

OS Saccharomyces cerevisiae.

```

XX FH Key Location/Qualifiers
XX FH Peptide 26..29
XX FT /note= "nuclear localisation signal"
XX FT Peptide 42..46
XX FT /note= "nuclear localisation signal"
XX FT Region 843..975
XX FT /note= "acidic region"
XX FT Protein 326..1089
XX FT /note= "C-terminal fragment, inhibits
XX FT nonsense-mediated mRNA decay"
XX PN W0920797-A1.
XX PD 29-APR-1999.
XX PF 21-OCT-1998; 98WO-US22365.
XX PR 21-OCT-1997; 97US-0955472.
XX (UYMA-) UNIV MASSACHUSETTS.
XX PI He F, Jacobson AS;
XX DR WPI; 1999-302756/25.
XX DR N-PSDB; AAX25601.
XX PT mRNA decay pathway genes and protein useful for treating breast
XX cancer and cystic fibrosis
XX PS
XX Example 3; Fig 1A-C; 116pp; English.
XX
XX This sequence represents Nmd2p, the protein encoded by the newly
XX discovered NMD2 gene (see AAX25601) of Saccharomyces cerevisiae.
XX NMD2 was named after its role in the nonsense-mediated mRNA decay
XX pathway. The Nmd2p protein binds to Upf1p, another protein in the
XX decay pathway. A C-terminal fragment of Nmd2p also binds to Upf1p
XX and, when overexpressed in a host cell, inhibits the function of
XX Upf1p, thereby inhibiting the nonsense-mediated mRNA decay pathway.
XX The invention also relates to methods of inhibiting the
XX nonsense-mediated mRNA decay pathway to stabilise mRNA transcripts
XX containing a nonsense codon which normally would cause an increased
XX translation from the transcript. Such stabilisation allows increased
XX production of a recombinant protein. The invention also relates to
XX methods of identifying molecules that inhibit the nonsense-mediated
XX mRNA decay pathway, and the use of such molecules (e.g. the
XX C-terminal fragment of Nmd2p or an antisense molecule) for
XX treatment of disorders associated with nonsense mutations, such
XX as breast cancer, polycystic kidney disease I and II, Niemann-Pick
XX disease, adenomatous polyposis coli, cystic fibrosis, Fanconi's
XX anaemia, haemophilia, hypercholesterolaemia, neurofibromatosis,
XX ornithine transcarbamylase deficiency, retinoblastoma, glycogen
XX storage disease, McArdle disease, cancer, Tay-Sachs disease, Cowden
XX disease, Wilson disease or beta-thalassemia (claimed).
XX
XX Sequence 1089 AA;
XX
XX Query Match 2.2%; Score 10; DB 20; Length 1089;
XX Best Local Similarity 100.0%; Pred. No. 1.9;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 437 DDDDDDDGEE 446
XX |||||
XX Db 902 DDDDDDDGEE 911

```

## RESULT 22

```

AA198055
ID AA198055 standard; Protein; 1089 AA.
XX AC AA198055;
XX
XX

```

```

DT 07-SEP-2000 (first entry)
XX XX
XX Yeast Nmd2p.
XX
XX Yeast; nonsense-mediated mRNA decay pathway; NMD2; UPF3; UPF1;
XX upf1p; upf3p; gene therapy; beta-thalassemia; cancer;
XX polycystic kidney disease; Duchenne muscular dystrophy; Fanconi anaemia;
XX haemophilia; hypercholesterolaemia; neurofibromatosis; Tay-Sachs disease;
XX glycogen storage disease; cystic fibrosis; adenomatous polyposis coli;
XX Cowden disease; Maple syrup urine disease; Wilson disease;
XX Niemann-Pick disease; Turcot syndrome; McArdle disease;
XX ornithine transcarbamylase deficiency; antisense therapy; ribozyme.
XX
XX Saccharomyces cerevisiae.
XX OS
XX FH Key Location/Qualifiers
XX FH Peptide 326..1089
XX FT /label= Carboxy terminal
XX FT /note= "Fragment is described in AA198056"
XX
XX US6071700-A.
XX PN
XX XX 06-JUN-2000.
XX PD
XX XX 21-OCT-1998; 98US-0177431.
XX PF
XX XX 20-JAN-1995; 95US-0375300.
XX PR 21-OCT-1997; 97US-0955472.
XX PA (UYMA-) UNIV MASSACHUSETTS.
XX XX
XX PI Jacobson AS, He F;
XX XX
XX WPI; 2000-422078/36.
XX DR N-PSDB; AAA39450.
XX
XX Testing for compounds able to modulate the nonsense mediated mRNA decay
XX pathway for increasing endogenous protein production by incubating a
XX candidate compound with a cell having a gene with a nonsense mutation
XX and a gene without mutation
XX
XX Disclosure; Fig 1; 57pp; English.
XX
XX The nonsense-mediated mRNA decay pathway has an important role in the
XX degradation of mRNA transcripts that contain a stop codon. The products
XX of the UPF1 and UPF3 genes (proteins Upf1p and Upf3p, respectively) are
XX essential components of this pathway. Mutations in these genes stabilise
XX mRNA containing premature stop codons. The present sequence is the NMD2
XX gene protein from Saccharomyces cerevisiae: Nmd2p. The carboxy terminal
XX of this protein binds to Upf1p to inhibit its function. Inhibition of the
XX nonsense-mediated mRNA decay pathway is a useful means of treating
XX disorders caused by the presence of nonsense mutations, e.g. treating
XX cancer, polycystic kidney diseases, Niemann-Pick disease, adenomatous
XX polyposis coli, cystic fibrosis, Fanconi's anaemia, haemophilia,
XX hypercholesterolaemia, neurofibromatosis, ornithine transcarbamylase
XX deficiency, retinoblastoma, glycogen storage disease, McArdle disease,
XX colorectal cancer, Tay-Sachs disease, Cowden disease, Wilson disease,
XX polycystic kidney disease, Duchenne muscular dystrophy, adrenocortical
XX carcinoma, Li-Fraumeni syndrome, Maple syrup urine disease, Turcot
XX syndrome or beta-thalassemia. Antisense mRNA molecules complementary to
XX the NMD2 gene may be used for antisense therapy of the above mentioned
XX disorders since the antisense molecules would inhibit translation of the
XX NMD2 mRNA. Additionally, the NMD2 gene may be manipulated for ribozyme
XX catalytic cleavage of the above mentioned disorders: ribozymes designed to
XX catalytically cleave nonsense-mediated mRNA decay pathway mRNAs e.g. NMD2
XX mRNA, can also be used to prevent translation and therefore expression of
XX these mRNAs.
XX
XX Sequence 1089 AA;
XX
XX Query Match 2.2%; Score 10; DB 21; Length 1089;
XX Best Local Similarity 100.0%; Pred. No. 1.9;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 437 DDDDDDDGEE 446  
Db 902 DDDDDDDGEE 911  
RESULT 23  
AAG21315  
ID AAG21315 standard; Protein: 133 AA.  
AC AAG21315;  
XX 17-OCT-2000 (first entry)  
DT  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23829.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0128845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145921.  
PR 28-JUL-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.

```

PR 13-AUG-1999; 99US-0146884.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149722.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151060.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158269.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 2.0%; Score 9; DB 21; Length 133;  
 Best Local Similarity 100.0%; Pred.No. 2.7; Mismatches 0; Indels 0; Gaps 0;

Matches \* 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 434 SEGDDDDDD 442  
 Db 47 SEGDDDDDD 55

RESULT 24

```

AAG21314
ID AAG21314 standard; Protein; 156 AA.
XX
AC AAG21314;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23828.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139459.

```

PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139751.  
PR 21-JUN-1999; 99US-0139817.  
PR 21-JUN-1999; 99US-0139818.  
PR 22-JUN-1999; 99US-0139819.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 24-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142927.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 20-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145087.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 22-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 26-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147182.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 13-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 13-OCT-1999; 99US-0159325.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 18-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 21-OCT-1999; 99US-0160880.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 25-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 26-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 2.0%; Score 9; DB 21; Length 156;  
Best Local Similarity 100.0%; Pred. No. 3.1;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 434 SEGD DDD 442  
DB 70 SEGD DDD 78

RESULT 25

AA07429 standard; Protein; 201 AA.

AA07429;

AC 17-OCT-2000 (first entry)  
XX  
DT  
XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 4579.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX Arabidopsis thaliana.  
PN EP1033405-A2.  
XX  
XX  
PD 06-SEP-2000.  
XX  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 16-APR-1999; 99US-0128714.  
PR 19-APR-1999; 99US-0129845.  
PR 21-APR-1999; 99US-0130077.  
PR 23-APR-1999; 99US-0130449.  
PR 28-APR-1999; 99US-0130510.  
PR 30-APR-1999; 99US-0130891.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 07-MAY-1999; 99US-0132486.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134768.  
PR 20-MAY-1999; 99US-0134941.  
PR 21-MAY-1999; 99US-0135124.  
PR 24-MAY-1999; 99US-0135353.  
PR 25-MAY-1999; 99US-0135629.  
PR 27-MAY-1999; 99US-0136021.  
PR 28-MAY-1999; 99US-0136392.  
PR 03-JUN-1999; 99US-0136782.  
PR 04-JUN-1999; 99US-0137528.  
PR 07-JUN-1999; 99US-0137502.  
PR 08-JUN-1999; 99US-0137722.  
PR 10-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140891.  
PR 30-JUN-1999; 99US-0140991.  
PR 01-JUL-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 12-JUL-1999; 99US-0142920.  
PR 13-JUL-1999; 99US-0142977.  
PR 14-JUL-1999; 99US-0143542.  
PR 15-JUL-1999; 99US-0143624.  
PR 16-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 19-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145218.  
PR 25-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145813.  
PR 27-JUL-1999; 99US-0145918.  
PR 28-JUL-1999; 99US-0145919.  
PR 02-AUG-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 09-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149928.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0149932.  
PR 23-AUG-1999; 99US-0149930.  
PR 26-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.

PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158233.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0158293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160815.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 25-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 26-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 28-OCT-1999; 99US-0162142.  
PR 29-OCT-1999; 99US-0162142.

Query Match 2.0%; Score 9; DB 21; Length 201;  
Best Local Similarity 100.0%; Pred. No. 3.8; 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0;

QY 434 SEGDGDDDD 442  
DB 47 SEGDGDDDD 55

RESULT 26.  
AAG07428  
ID AAG07428 standard; Protein; 224 AA.

XX 17-OCT-2000 (first entry)  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 4578.  
XX Protein identification; signal transduction pathway; metabolic pathway;  
XX Hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.

XX PN EPI033405-A2.  
XX PD 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
XX 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
XX 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
XX 28-APR-1999; 99US-0130891.  
XX 30-APR-1999; 99US-0131449.  
XX 30-APR-1999; 99US-0132048.  
XX 04-MAY-1999; 99US-0132407.  
XX 05-MAY-1999; 99US-0132485.  
XX 06-MAY-1999; 99US-0132486.  
XX 07-MAY-1999; 99US-0132487.  
XX 11-MAY-1999; 99US-0132863.  
XX 11-MAY-1999; 99US-0134256.  
XX 14-MAY-1999; 99US-0134218.  
XX 14-MAY-1999; 99US-0134219.  
XX 14-MAY-1999; 99US-0134221.  
XX 14-MAY-1999; 99US-0134370.  
XX 18-MAY-1999; 99US-0134768.  
XX 19-MAY-1999; 99US-0134941.  
XX 20-MAY-1999; 99US-0135124.  
XX 21-MAY-1999; 99US-0135533.  
XX 24-MAY-1999; 99US-0135629.  
XX 25-MAY-1999; 99US-0136021.  
XX 27-MAY-1999; 99US-0136392.  
XX 28-MAY-1999; 99US-0136782.  
XX 01-JUN-1999; 99US-0137222.  
XX 03-JUN-1999; 99US-0137528.  
XX 04-JUN-1999; 99US-0137502.  
XX 07-JUN-1999; 99US-0137724.  
XX 08-JUN-1999; 99US-0138094.  
XX 10-JUN-1999; 99US-0138540.  
XX 10-JUN-1999; 99US-0138847.  
XX 14-JUN-1999; 99US-0139119.  
XX 16-JUN-1999; 99US-0139452.  
XX 17-JUN-1999; 99US-0139453.  
XX 17-JUN-1999; 99US-0139452.  
XX 18-JUN-1999; 99US-0139454.  
XX 18-JUN-1999; 99US-0139455.  
XX 18-JUN-1999; 99US-0139456.  
XX 18-JUN-1999; 99US-0139457.  
XX 18-JUN-1999; 99US-0139458.  
XX 18-JUN-1999; 99US-0139459.  
XX 18-JUN-1999; 99US-0139460.  
XX 18-JUN-1999; 99US-0139461.  
XX 18-JUN-1999; 99US-0139462.  
XX 18-JUN-1999; 99US-0139463.  
XX 18-JUN-1999; 99US-0139464.  
XX 18-JUN-1999; 99US-0139465.  
XX 21-JUN-1999; 99US-0139817.  
XX 22-JUN-1999; 99US-0139889.  
XX 23-JUN-1999; 99US-0140353.  
XX 23-JUN-1999; 99US-0140354.  
XX 24-JUN-1999; 99US-0140695.  
XX 28-JUN-1999; 99US-0140823.  
XX 29-JUN-1999; 99US-0140991.  
XX 30-JUN-1999; 99US-0141287.



```

PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142970.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144844.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147933.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149928.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.

```

```

PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match      2.0%; Score 9; DB 21; Length 224;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 434 SEGD DDD 442
DB 70 SEGD DDD 78

```

```

RESULT 27
AAR55371
ID AAR55371 standard; Protein; 502 AA.
XX
AC AAR55371;
XX
DT 20-JAN-1995 (first entry)
XX
DE Mouse Activin receptor-like kinase 1 (mALK-1).
XX
KW serine threonine kinases; activin receptors; Act-R; superfamily;
XX transforming growth factor; TGF; diagnostics; detection; therapy;
KW rheumatoid arthritis; glomerular nephritis; fibrosis.
XX
OS Mus musculus.
XX
PN WO9411502-A.
XX
PD 26-MAY-1994.
XX
PF 17-NOV-1993; 93WO-GB02367.
XX

```



	CC	hydrophobic single transmembrane domain of 23-25 residues, an
	CC	intracellular serine/threonine kinase domain of approximately 300
	CC	residues and a short serine/threonine-rich tail. The present sequence
	CC	is that of MIRS3 which is believed to be a monomeric isoform of the
	CC	rat tubulin receptor and/or BMP receptor.
	CC	Isolated DNA which encodes the MIRS4 sequence or which is able to
	CC	hybridise to such DNA under stringent conditions is claimed.
	XX	
SQ	Sequence	505 AA;
	Query Match	2.0%; Score 9; DB 17; Length 505;
Beech Local Similarity	100.0%;	Pred. No. 8.5;
Matches	9; Conservative	0; Mismatches 0; Indels 0; Gaps 0
OY	131 HRDLKSRNV 139	
Dd	330 HRDLKSRNV 338	
RESULT 31		
ID	ABP30294	ABP30294 standard; Protein; 648 AA.
XX		
AC	ABP30294;	
XX		
DT	02-JUL-2002	(first entry)
XX		
DE	Streptococcus polypeptide SEQ ID NO 9764.	
KW	Streptococcus: GAS; GBS; group B streptococcus; Streptococcus agalactiae;	
KM	group A streptococcus; Streptococcus pyogenes; antibacterial;	
KN	antiinflammatory; infection; vaccine; meningitis; gene therapy.	
OS	Streptococcus agalactiae.	
PN	WO200234771-A2.	
PD	02-MAY-2002.	
PF	29-OCT-2001; 2001MO-GB04789.	
XX		
PR	27-OCT-2000; 2000GB-0026333.	
PR	24-NOV-2000; 2000GB-0028727.	
XX	07-MAR-2001; 2001GB-0005640.	
PA	(CHIR-) CHIRON SPA.	
PA	(GENO-) INST GENOMIC RES.	
PI	Telford J, Masignani V, Margarit Ros YI, Grandt G, Fraese C;	
PI	Tettelein H;	
DR	WPI; 2002-352536/38.	
DR	N-PSDB; ABN70925.	
PT	New Streptococcus protein for the treatment or prevention of infection	
PT	or disease caused by Streptococcus bacteria, such as meningitis, and	
PT	for detecting a compound that binds to the protein -	
XX		
PS	Claim 1; Page 4103; 4525pp; English.	
CC	The invention relates to a protein (ABP25413-ABP30895) from group B	
CC	streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS	
CC	(Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in	
CC	the specification. The proteins have antibacterial and anti-inflammatory	
CC	activity. (I), nucleic acids encoding (II), ABN66044-ABN71526 and	
CC	antibodies that bind (IV) are used in the manufacture of medicaments for	
CC	the treatment or prevention of infection or disease caused by	
CC	Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.	
CC	Nucleic acids encoding (I) are used to detect Streptococcus in a	
CC	biological sample. (I) is used to determine whether a compound binds to	
CC	(I). A composition comprising (I) or a nucleic acid encoding (I), may be	
CC	used as a vaccine or diagnostic composition. The disease caused by	
CC	Streptococcus	

CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC used in gene therapy. Antibodies to (I) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins.

XX Sequence 648 AA;

Query Match 2.0%; Score 9; DB 23; Length 648;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 EKEAEILSV 59  
 |||||  
 Db 303 EKEAEILSV 311

RESULT 32  
 ABP27599  
 ID ABP27599 standard; Protein; 652 AA.

XX ABP27599;

DT 02-JUL-2002 (first entry)

XX Streptococcus polypeptide SEQ ID NO 4374.

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

XX group A streptococcus; Streptococcus pyogenes; antibacterial;

XX antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX Streptococcus agalactiae.

XX WO200234771-A2.

XX 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB04789.

XX 27-OCT-2000; 2000GB-0026333.

XX 24-NOV-2000; 2000GB-0028727.

XX 07-MAR-2001; 2001GB-0005640.

XX (CHIR-) CHIRON SPA.

XX (GENO-) INST GENOMIC RES.

XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;

XX Tettelin H;

XX MPI: 2002-352536/38.

XX N-PSDB; ABN68230.

XX New Streptococcus protein for the treatment or prevention of infection

XX or disease caused by Streptococcus bacteria, such as meningitis, and

XX for detecting a compound that binds to the protein -

XX Claim 1; Page 3598; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B

XX streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS

XX (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in

XX the specification. The proteins have antibacterial and antiinflammatory

XX activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and

XX antibodies that bind (I) are used in the manufacture of medicaments for

XX the treatment or prevention of infection or disease caused by

XX Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

XX Nucleic acids encoding (I) are used to detect Streptococcus in a

XX biological sample. (I) is used to determine whether a compound binds to

XX (I). A composition comprising (I) or a nucleic acid encoding (I), may be

XX used as a vaccine or diagnostic composition. The disease caused by

XX Streptococcus that is prevented or treated may be meningitis. Nucleic

XX acid encoding (I) may be used to recombinantly produce (I) and may be

XX used in gene therapy. Antibodies to (I) are used for affinity

XX chromatography, immunoassays, and distinguishing/identifying

CC Streptococcus proteins.

XX Sequence 652 AA;

XX Query Match 2.0%; Score 9; DB 23; Length 652;

XX Best Local Similarity 100.0%; Pred. No. 11;

XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 EKEAEILSV 59  
 |||||  
 Db 307 EKEAEILSV 315

RESULT 33  
 ABP27600  
 ID ABP27600 standard; Protein; 652 AA.

XX ABP27600;

DT 02-JUL-2002 (first entry)

XX Streptococcus polypeptide SEQ ID NO 4376.

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

XX group A streptococcus; Streptococcus pyogenes; antibacterial;

XX antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX Streptococcus pyogenes.

XX WO200234771-A2.

XX 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB04789.

XX 27-OCT-2000; 2000GB-0026333.

XX 24-NOV-2000; 2000GB-0028727.

XX 07-MAR-2001; 2001GB-0005640.

XX (CHIR-) CHIRON SPA.

XX (GENO-) INST GENOMIC RES.

XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;

XX Tettelin H;

XX MPI: 2002-352536/38.

XX N-PSDB; ABN68231.

XX New Streptococcus protein for the treatment or prevention of infection

XX or disease caused by Streptococcus bacteria, such as meningitis, and

XX for detecting a compound that binds to the protein -

XX Claim 1; Page 3598; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B

XX streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS

XX (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in

XX the specification. The proteins have antibacterial and antiinflammatory

XX activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and

XX antibodies that bind (I) are used in the manufacture of medicaments for

XX the treatment or prevention of infection or disease caused by

XX Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

XX Nucleic acids encoding (I) are used to detect Streptococcus in a

XX biological sample. (I) is used to determine whether a compound binds to

XX (I). A composition comprising (I) or a nucleic acid encoding (I), may be

XX used as a vaccine or diagnostic composition. The disease caused by

XX Streptococcus that is prevented or treated may be meningitis. Nucleic

XX acid encoding (I) may be used to recombinantly produce (I) and may be

XX used in gene therapy. Antibodies to (I) are used for affinity

XX chromatography, immunoassays, and distinguishing/identifying

XX Streptococcus proteins.

XX Sequence 652 AA;

Query Match 2.0%; Score 9; DB 23; Length 652;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

51 EKEEILSV 59  
 |||||  
 307 EKEEILSV 315

RESULT 34  
 AA016259 standard; Protein; 656 AA.

AA016259;  
 26-MAR-2002 (first entry)  
 Human kinase PKIN-5 protein.

Human; kinase; PKIN-5; cancer; leukaemia; adenocarcinoma; osteoarthritis;  
 immune disorder; atherosclerosis; Crohn's disease; Hodgkin's disease;  
 Acquired Immune Deficiency Syndrome; AIDS; Addison's disease; anaemia;  
 allergy; asthma; adult respiratory distress syndrome; multiple sclerosis;  
 autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoporosis;  
 Good pasture's syndrome; Graves' disease; pancreatitis; psoriasis;  
 rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome;  
 hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris;  
 cardiovascular disease; hypertension; vasculitis; myocarditis; obesity;  
 congestive heart failure; ischaemic heart disease; lung tumour; gout;  
 fatty liver; Niemann-Pick's disease; gene therapy.

Homo sapiens.

Location/Qualifiers.  
 13..252  
 /label= Protein\_kinase\_domain  
 14..197  
 /label= Protein\_kinase\_domain  
 14..257  
 /note= "Eukaryotic protein kinase domain"  
 14..252  
 /label= Protein\_kinase\_domain  
 14..253  
 /label= Protein\_kinase\_domain

WO200196547-A2.  
 20-DEC-2001.  
 14-JUN-2001; 2001WO-US19444.  
 15-JUN-2000; 2000US-212073P.  
 23-JUN-2000; 2000US-213467P.  
 30-JUN-2000; 2000US-215651P.  
 07-JUL-2000; 2000US-21605P.  
 13-JUL-2000; 2000US-218372P.  
 25-AUG-2000; 2000US-228056P.

(INCY-) INCYTE GENOMICS INC.

Yue H, Lal P, Bandman O, Borowsky ML, Au-Young J, Lu Y,  
 Gandhi AR, Tribouley CM, Walla NK, Yao MG, Lu DM, Greenwald SR,  
 Rankum J, Griffin JA, Kearney L, Burford N, Nguyen DB, Tang YF,  
 Baughn MR, He A, Thornton M, Hatala A, Patterson C, Gururajan R,  
 Lo TP, Khan F, Recton SA, Azimzai Y, Pollock JL, Ding L,  
 Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH;  
 WPI; 2002-090207/12.  
 N-PSDB; AAD26452.

New polypeptides, useful for diagnosing, treating or preventing  
 disorders of growth and development, cardiovascular and lipid, and

PT diseases such as cancer, comprise human kinase polypeptides -  
 XX Claim 1; Page 143-144; 197pp; English.  
 PS  
 XX The invention relates to human kinase PKIN proteins and their  
 CC corresponding cDNAs. A composition containing PKIN agonist is useful for  
 CC treating a disease or condition associated with decreased expression of  
 CC PKIN and a composition comprising PKIN antagonist is useful for treating  
 CC a disease or condition associated with overexpression of PKIN. The  
 CC disorders include cancer (leukaemia, adenocarcinoma, lymphoma, melanoma,  
 CC myeloma, sarcoma, teratocarcinoma, Hodgkin's disease), immune disorder  
 CC (Acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease,  
 CC atherosclerosis, anaemia, allergies, adult respiratory distress syndrome,  
 CC autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes  
 CC mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease,  
 CC osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome,  
 CC rheumatoid arthritis, Sjogren's syndrome, uveitis, ulcerative colitis,  
 CC bacterial, parasitic, fungal, viral, protozoal and helminthic infections)  
 CC growth and development disorders (arteriosclerosis, cirrhosis, hepatitis,  
 CC Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts); cardio  
 CC vascular disease (arteriovenous fistula, hypertension, vasculitis,  
 CC aneurysms, congestive heart failure, angina pectoris, myocarditis,  
 CC ischaemic heart disease, chronic bronchitis, lung tumours); lipid  
 CC disorder (fatty liver, Fabry's disease, Niemann-Pick's disease,  
 CC hypercholesterolaemia, obesity). PKIN DNA is useful for assessing  
 CC toxicity of a test compound and in gene therapy. The present sequence  
 CC is human PKIN-5 protein.

Sequence 656 AA;  
 25 GSFGSVYRA 33  
 |||||  
 17 GSFGSVYRA 25

Query Match 2.0%; Score 9; DB 23; Length 656;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

25 GSFGSVYRA 33  
 |||||  
 17 GSFGSVYRA 25

RESULT 35  
 AA003554  
 ID AA003554 standard; Protein; 909 AA.

AA003554;  
 12-SEP-2001 (first entry)  
 Human protein kinase #54.

Human; protein kinase; PK; cancer; cardiovascular disease;  
 metabolic disorder; immune related disease; neurological disorder;  
 neurodegenerative disorder; inflammatory disorder; infectious disease;  
 reproductive disorder.

Homo sapiens.

WO200138503-A2.  
 31-MAY-2001.  
 22-NOV-2000; 2000WO-US32085.  
 24-NOV-1999; 99US-0167482.  
 (SUGB-) SUGEN INC.

Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R,  
 Flanagan P, Clary D;  
 WPI; 2001-343950/36.  
 N-PSDB; AAS06754.

Nucleic acids encoding human kinase polypeptides, useful for preventing

PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and  
 PT neuronal-associated diseases, and microbial infections -

XX Claim 7; Figure 2; 433bp; English.

XX AAU03501-AAU03557 represent novel human protein kinases #1-57. The  
 CC novel protein kinases have been identified as members of the tyrosine  
 CC or serine/threonine kinase (PTK and STK) families. The polynucleotides  
 CC encoding protein kinases and treatment of diseases associated with  
 CC preventing, diagnosis and treatment of diseases associated with  
 CC inappropriate kinase expression. For example, they may be used to treat  
 CC cancers (especially cancers of hematopoietic origin), cardiovascular  
 CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),  
 CC immune related diseases (e.g. rheumatoid arthritis), neurological  
 CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.  
 CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious  
 CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).  
 CC Additionally, polynucleotides encoding protein kinases may be  
 CC used for gene therapy and as DNA probes in diagnostic assays.  
 CC The protein kinase polypeptides may be used as antigens in the production  
 CC of antibodies against the protein kinases and in assays to identify  
 CC modulators of protein kinase expression and activity.

XX Sequence 909 AA;

Query Match 2.0%; Score 9; DB 22; Length 909;

Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 GSFGSYTRA 33  
 |||||

DB 403 GSFGSYTRA 411

RESULT 36

AAK55694  
 ID AAK55694 standard; Protein; 2391 AA.

XX AAK55694;

DT 06-DEC-1994 (first entry)

XX Cardamoyl-phosphate-synthetase II.

KW Cardamoyl-phosphate-synthetase II; CPSII; pCPSII gene;

KM malaria.

XX Plasmodium falciparum.

CS Plasmodium falciparum.

Key Location/Qualifiers  
 FT Domain 1..690 /note= "glutamine-amidotransferase domain"

FT Domain 1..270 /note= "structural subdomain"

FT Peptide 271..482 /note= "insert sequence"

FT Domain 483..690 /note= "glutaminase subdomain"

FT Domain 691..2391 /note= "Cardamoyl-phosphate-synthase domain"

FT Peptide 1255..1857 /note= "ATP binding subdomain CPSa"

FT Domain 1858..2391 /note= "insert sequence"

PN WO9412643-A.

PD 09-JUN-1994.

PF 02-DEC-1993; 93WO-AU00617.

PR 03-DEC-1992; 92AU-0006206.  
 PR 16-DEC-1992; 92AU-0006380.  
 XX (UNIX) UNISEARCH LTD.

PI Flores MV, Oesullivan WJ, Stewart TS;

DR WPI, 1994-200271/24.

XX N-PSDB; AAQ62924.

PT Nucleic acid encoding cardamoyl phosphate synthetase II -  
 PT isolated from Plasmodium falciparum, used to develop prods. for  
 PT the treatment of malaria.

PS Disclosure; Page 6-16; 31pp; English.

CC The cDNA sequence encoding the cardamoyl-phosphate-transferase II  
 CC (CPSII) of Plasmodium falciparum was determined. The cDNA encodes  
 CC a protein that includes 2 insert sequences not found in other CPSII  
 CC proteins. The first separates the putative structural subdomain and  
 CC the glutaminase subdomain of the glutamine-amidotransferase subunit  
 CC of CPSII, while the second separates 2 ATP binding subdomains of the  
 CC CPSII subunit, CPSa and CPSb.

XX Sequence 2391 AA;

Query Match 2.0%; Score 9; DB 15; Length 2391;

Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDGE 445  
 |||||

DB 1786 DDDDDDDGE 1794

RESULT 37

AAK83387  
 ID AAK83387 standard; peptide; 20 AA.

XX AAK83387;

DT 26-MAR-2002 (first entry)

XX Lunasin fragment #8.

KW Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;  
 KM prostate; liver; kidney; therapy; soybean.

XX Glycine max.

OS WO200172784-A2.

PD 04-OCT-2001.

XX 23-MAR-2001; 2001WO-US09453.

PR 24-MAR-2000; 2000US-0534705.

PA (FLLG-) FLLGEN BIOSCIENCES INC.

PI Galvez AF;

DR WPI, 2001-648438/74.

PT Treatment or prevention of cancer by administering a lunasin peptide  
 PT which binds to non-acetylated histones H3 and H4 and prevents their  
 PT acetylation in hypocyclated repressed chromatin -

PS Claim 7; Page 42; 49pp; English.

CC This sequence represents a lunasin peptide used in the method of the  
 CC invention. The method is for the treatment or prevention of cancer,  
 CC comprising administration of a lunasin peptide or its active fragment or

CC analog. The invention is used to treat or prevent cancer, particularly  
 CC in a human. The cancers that can be treated and prevented include  
 CC those of the colon, upper gastrointestinal tract, breast, prostate,  
 CC liver, kidney or any other internal organs or tissues.  
 XX  
 SQ Sequence 20 AA;

Query Match 1.8%; Score 8; DB 22; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 4.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 436 GDDDDDD 443  
 |||||  
 DB 13 GDDDDDD 20

RESULT 38  
 AAB62621  
 ID AAB62621 standard; peptide; 20 AA.  
 AAB62621;

DT 23-JUL-2001 (first entry)

DE Soybean lunasin peptide variant (residues 22-41).

KM Lunasin; cancer; soybean; variant; carcinogen; viral; oncogene.

OS Glycine max.

PN WO200134808-A2.

PD 17-MAY-2001.

PF 12-NOV-2000; 2000WO-US31211.

PR 12-NOV-1999; 99US-0165334.

(FIG-) FILGEN BIOSCIENCES INC.

Galvez AF;

WPI; 2001-343605/36.

Large scale production of lunasin, a cancer preventive peptide from  
 soybean, by preparing lunasin gene constructs using Pichia expression  
 vectors, optimizing lunasin expression parameters and purifying lunasin

Claim 7; Page 62; 67pp; English.

The invention provides a method of producing lunasin, a cancer preventive  
 peptide from soybean, or its active variant by recombinant DNA technology  
 in large quantities. The method comprises preparing lunasin gene  
 constructs or its variants using protein expression vectors, optimizing  
 lunasin expression parameters for large-scale production and performing  
 a series of isolation and purification to obtain large quantities of  
 biologically active recombinant lunasin peptide. The lunasin peptide  
 exhibits inhibitory effects against malignant transformation of cells  
 induced by chemical carcinogens and viral oncogenes. Sequences  
 AAB62614-24 represent soybean lunasin variants.

Sequence 20 AA;

Query Match 1.8%; Score 8; DB 22; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 4.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
 |||||  
 DB 13 GDDDDDD 20

RESULT 39  
 AAB83386  
 ID AAB83386 standard; peptide; 21 AA.  
 AAB83386;

DT 26-MAR-2002 (first entry)  
 DE Lunasin fragment #7.

KM Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;  
 KW prostate; liver; kidney; therapy; soybean.

OS Glycine max.

PN WO200172784-A2.

PD 04-OCT-2001.

PF 23-MAR-2001; 2001WO-US09453.

PR 24-MAR-2000; 2000US-0534705.

(FIG-) FILGEN BIOSCIENCES INC.

Galvez AF;

WPI; 2001-648438/74.

Treatment or prevention of cancer by administering a lunasin peptide  
 which binds to non-acetylated histones H3 and H4 and prevents their  
 acetylation in hypoacetylated repressed chromatin

Claim 7; Page 42; 49pp; English.

This sequence represents a lunasin peptide used in the method of the  
 invention. The method is for the treatment or prevention of cancer,  
 comprising administration of a lunasin peptide or its active fragment or  
 analog. The invention is used to treat or prevent cancer, particularly  
 those of the colon, upper gastrointestinal tract, breast, prostate,  
 liver, kidney or any other internal organs or tissues.

Sequence 21 AA;

Query Match 1.8%; Score 8; DB 22; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 4.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
 |||||  
 DB 13 GDDDDDD 20

RESULT 40  
 AAB83393  
 ID AAB83393 standard; peptide; 21 AA.  
 AAB83393;

DT 26-MAR-2002 (first entry)

DE Lunasin related peptide #2.

KM Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;  
 KW prostate; liver; kidney; therapy; soybean.

OS Glycine max.

PN WO200172784-A2.

PD 04-OCT-2001.

PF 23-MAR-2001; 2001WO-US09453.  
XX  
XX 24-MAR-2000; 2000US-0534705.  
PR  
XX (FTLG-) FILGEN BIOSCIENCES INC.  
PA  
XX Galvez AF;  
PI  
XX WPI; 2001-648438/74.  
DR  
XX Treatment or prevention of cancer by administering a lunasin peptide  
PT which binds to non-acetylated histones H3 and H4 and prevents their  
PT acetylation in hypoacetylated repressed chromatin  
PS  
XX Disclosure; Page 46; 49pp; English.  
XX This sequence represents a lunasin related peptide used in the method of  
CC the invention. The method is for the treatment or prevention of cancer,  
CC comprising administration of a lunasin peptide or its active fragment or  
CC analog. The invention is used to treat or prevent cancer, particularly  
in a human. The cancers that can be treated and prevented include  
those of the colon, upper gastrointestinal tract, breast, prostate,  
CC liver, kidney or any other internal organs or tissues.  
CC  
SQ Sequence 21 AA;  
Query Match 1.8%; Score 8; DB 22; Length 21;  
Best Local Similarity 100.0%; Pred. No. 4.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 436 GDDDDDD 443  
DB 12 GDDDDDD 19  
RESULT 41  
AAB62620  
ID AAB62620 standard; peptide; 21 AA.  
XX  
XX AAB62620;  
AC  
XX 23-UTL-2001 (first entry)  
DT  
XX Soybean lunasin peptide variant (residues 22-42).  
DE  
XX Lunasin; cancer; soybean; variant; carcinogen; viral; oncogene.  
KM  
XX Glycine max.  
AS  
XX WO200134808-A2.  
XX  
XX 17-MAY-2001.  
PD  
XX 12-NOV-2000; 2000WO-US31211.  
PF  
XX 12-NOV-1999; 99US-0165334.  
PR  
XX (FTLG-) FILGEN BIOSCIENCES INC.  
PA  
XX Galvez AF;  
PI  
XX WPI; 2001-343605/36.  
DR  
XX Large scale production of lunasin, a cancer preventive peptide from  
PT soybean, by preparing lunasin gene constructs using Pichia expression  
PT vectors, optimizing lunasin expression parameters and purifying lunasin  
PT  
XX  
XX Claim 7; Page 62; 67pp; English.  
PS  
XX The invention provides a method of producing lunasin, a cancer preventive  
CC peptide from soybean, or its active variant by recombinant DNA technology  
CC in large quantities. The method comprises preparing lunasin gene

CC constructs or its variants using protein expression vectors, optimizing  
CC lunasin expression parameters for large-scale production and performing  
CC a series of isolation and purification to obtain large quantities of  
CC biologically active recombinant lunasin peptide. The lunasin peptide  
CC exhibits inhibitory effects against malignant transformation of cells  
CC induced by chemical carcinogens and viral oncogenes. Sequences  
CC AAB62614-24 represent soybean lunasin variants.  
CC  
SQ Sequence 21 AA;  
Query Match 1.8%; Score 8; DB 22; Length 21;  
Best Local Similarity 100.0%; Pred. No. 4.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 436 GDDDDDD 443  
DB 13 GDDDDDD 20  
RESULT 42  
AAB83385  
ID AAB83385 standard; peptide; 22 AA.  
XX  
XX AAB83385;  
AC  
XX 26-MAR-2002 (first entry)  
DT  
XX Lunasin fragment #6.  
DE  
XX Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;  
XX prostate; liver; kidney; therapy; soybean.  
KM  
XX Glycine max.  
OS  
XX WO200172784-A2.  
PN  
XX 04-OCT-2001.  
PD  
XX 23-MAR-2001; 2001WO-US09453.  
PF  
XX 24-MAR-2000; 2000US-0534705.  
PR  
XX (FTLG-) FILGEN BIOSCIENCES INC.  
PA  
XX Galvez AF;  
PI  
XX WPI; 2001-648438/74.  
DR  
XX Treatment or prevention of cancer by administering a lunasin peptide  
PT which binds to non-acetylated histones H3 and H4 and prevents their  
PT acetylation in hypoacetylated repressed chromatin  
PS  
XX Claim 7; Page 42; 49pp; English.  
XX This sequence represents a lunasin peptide used in the method of the  
CC invention. The method is for the treatment or prevention of cancer,  
CC comprising administration of a lunasin peptide or its active fragment or  
CC analog. The invention is used to treat or prevent cancer, particularly  
in a human. The cancers that can be treated and prevented include  
those of the colon, upper gastrointestinal tract, breast, prostate,  
CC liver, kidney or any other internal organs or tissues.  
CC  
SQ Sequence 22 AA;  
Query Match 1.8%; Score 8; DB 22; Length 22;  
Best Local Similarity 100.0%; Pred. No. 5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 436 GDDDDDD 443  
DB 13 GDDDDDD 20



RESULT 43  
ABB30607  
ID ABB30607 standard; Peptide; 22 AA.  
XX  
AC ABB30607;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Peptide #3258 encoded by breast cell single exon nucleic acid probe.  
XX  
KW Human; microarray; single exon probe; gene expression; breast;  
XX disease; cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157271-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00662.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
XX 26-MAY-2000; 2000US-0207456.  
XX 30-JUN-2000; 2000US-0608408.  
XX 21-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-0234687.  
XX 27-SEP-2000; 2000US-0236359.  
XX 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-496933/54.  
XX  
PT New spatially-addressable set of single exon nucleic acid probes,  
XX useful for measuring gene expression in sample derived from human  
XX breast, comprises number of single exon nucleic acid probes  
XX  
PS Claim 27; SEQ ID NO 13575; 327bp + sequence listing; English.  
XX  
XX The invention relates to a spatially-addressable set of single exon  
XX nucleic acid probes for measuring gene expression in a sample derived  
XX from human breast and BT 474 cells. The method involves contacting  
XX the probes with a collection of detectably labelled nucleic acids  
XX derived from mRNA of human breast, and then measuring the label  
XX bound to each probe of the microarray. The probes are useful for  
XX verifying the expression of the microarray. The probes are useful for  
XX encoding proteins. They are useful for gene discovery, and for  
XX determining predisposition and/or prognosing breast disease. Gene  
XX expression analysis is useful for assessing the toxicity of chemical  
XX agents on cells. The microarray of this invention presents a far greater  
XX diversity of probes for measuring gene expression, with far less bias  
XX than expressed sequence tag microarrays. The method is suitable for  
XX rapid production of functional information from genomic sequence. The  
XX present sequence is a peptide encoded by a single exon nucleic acid  
XX probe of the invention.  
XX Note: The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 22 AA;  
XX  
XX Query Match 1.8%; Score 8; DB 22; Length 22;  
XX Best Local Similarity 100.0%; Pred. No. 5;  
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 437 DDDDDDDG 444  
XX |||||  
XX  
DB 7 DDDDDDDG 14  
XX  
RESULT 44

ABB35771  
ID ABB35771 standard; Peptide; 22 AA.  
XX  
AC ABB35771;  
XX  
DT 04-FEB-2002 (first entry)  
XX  
DE Peptide #3277 encoded by human foetal liver single exon probe.  
XX  
KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
XX  
OS Homo sapiens.  
XX  
PN WO200157277-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00669.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
XX 26-MAY-2000; 2000US-0207456.  
XX 30-JUN-2000; 2000US-0608408.  
XX 03-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-0234687.  
XX 27-SEP-2000; 2000US-0236359.  
XX 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483447/52.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human foetal liver  
XX  
PS Claim 27; SEQ ID NO 28406; 639bp + sequence listing; English.  
XX  
XX The invention relates to a single exon nucleic acid probe for  
XX measuring human gene expression in a sample derived from human foetal  
XX liver. The single exon nucleic acid probes may be used for predicting,  
XX measuring and displaying gene expression in samples derived from human  
XX foetal liver. The present sequence is a peptide encoded by a single exon  
XX nucleic acid probe of the invention.  
XX Note: The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 22 AA;  
XX  
XX Query Match 1.8%; Score 8; DB 22; Length 22;  
XX Best Local Similarity 100.0%; Pred. No. 5;  
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 437 DDDDDDDG 444  
XX |||||  
XX  
DB 7 DDDDDDDG 14  
XX  
RESULT 45  
ABB21195  
ID ABB21195 standard; Protein; 22 AA.  
XX  
AC ABB21195;  
XX  
DT 23-JAN-2002 (first entry)  
XX  
DE Protein #3194 encoded by probe for measuring heart cell gene expression.  
XX  
KW Human; gene expression; heart; microarray; vascular system;  
XX cardiovascular disease; hypertension; cardiac arrhythmia;  
XX congenital heart disease.  
XX

```

OS Homo sapiens.
XX
XX WO200157274-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00666.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX
XX Claim 15; SEQ ID No 22965; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart (see
XX ABA2155-AA41305). The present sequence is a protein encoded by one such
XX probe. The probes may be used for predicting, measuring and displaying
XX gene expression in samples derived from the human heart via microarrays.
XX By measuring gene expression, the probes are useful for predicting,
XX diagnosing, grading, staging, monitoring and prognosing diseases of the
XX human heart and vascular system e.g. cardiovascular disease,
XX hypertension, cardiac arrhythmias and congenital heart disease.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 22 AA;
XX
XX Query Match 1.8%; Score 8; DB 22; Length 22;
XX Best Local Similarity 100.0%; Pred. No. 5;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 437 DDDDDDDG 444
XX 7 DDDDDDDG 14
XX
XX RESULT 46
XX ID AAM56580
XX AAM56580 standard; Protein; 22 AA.
XX
XX AC AAM56580;
XX
XX DT 05-NOV-2001 (first entry)
XX
XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 28685.
XX
XX KW Human; brain expressed exon; gene expression analysis; probe;
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer.
XX
XX OS Homo sapiens.
XX
XX PN WO200157275-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US00667.
XX

```

```

PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4; SEQ ID NO: 28685; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention.
XX
XX Sequence 22 AA;
XX
XX Query Match 1.8%; Score 8; DB 22; Length 22;
XX Best Local Similarity 100.0%; Pred. No. 5;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 437 DDDDDDDG 444
XX 7 DDDDDDDG 14
XX
XX RESULT 47
XX ID AAM68960
XX AAM68960 standard; Protein; 22 AA.
XX
XX AC AAM68960;
XX
XX DT 06-NOV-2001 (first entry)
XX
XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 29266.
XX
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukemia; lymphoma; myeloma.
XX
XX OS Homo sapiens.
XX
XX PN WO200157276-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US00668.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488900/53.
XX

```

PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human bone marrow -  
 XX  
 PS Example 4; SEQ ID NO: 29266; 658bp + Sequence Listing; English.  
 CC  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention.  
 XX  
 SQ Sequence 22 AA;

Query Match 1.8%; Score 8; DB 22; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 437 DDDDDDDG 444  
 D 7 DDDDDDDG 14

## RESULT 48

AA16787  
 ID AA16787 standard; Protein; 22 AA.

AC AA16787;  
 XX

DT 12-OCT-2001 (first entry)  
 XX

DE Peptide #3221 encoded by probe for measuring cervical gene expression.  
 XX

KW Probe; human; microarray; gene expression; cervical epithelial cell;  
 XX

OS Homo sapiens.  
 XX

PN WO200157278-A2.  
 XX

PD 09-AUG-2001.  
 XX

PF 30-JAN-2001; 2001WO-US00670.  
 XX

PR 04-FEB-2000; 2000US-0180312.  
 XX

PR 26-MAY-2000; 2000US-0207456.  
 XX

PR 30-JUN-2000; 2000US-0608408.  
 XX

PR 03-AUG-2000; 2000US-0632366.  
 XX

PR 21-SEP-2000; 2000US-0234687.  
 XX

PR 27-SEP-2000; 2000US-0236359.  
 XX

PR 04-OCT-2000; 2000GB-0024263.  
 XX

PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX

DR WPI; 2001-488901/53.  
 XX

PT Human genome-derived single exon nucleic acid probes useful for  
 XX analyzing gene expression in human cervical epithelial cells -

PS Claim 27; SEQ ID No 21613; 487bp; English.  
 XX

XX The present invention relates to human single exon nucleic acid probes  
 CC (SENPs; see A110066-A128459). The present sequence is a peptide encoded  
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 22 AA;

Query Match 1.8%; Score 8; DB 22; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 437 DDDDDDDG 444  
 Db 7 DDDDDDDG 14

## RESULT 49

AA129271  
 ID AA129271 standard; Protein; 22 AA.

AC AA129271;  
 XX

DT 17-OCT-2001 (first entry)  
 XX

DE Peptide #3308 encoded by probe for measuring placental gene expression.  
 XX

KW Probe; microarray; human; placenta; antenatal diagnosis;  
 XX genetic disorder.

OS Homo sapiens.  
 XX

PN WO200157272-A2.  
 XX

PD 09-AUG-2001.  
 XX

PF 30-JAN-2001; 2001WO-US00663.  
 XX

PR 04-FEB-2000; 2000US-0180312.  
 XX

PR 26-MAY-2000; 2000US-0207456.  
 XX

PR 30-JUN-2000; 2000US-0608408.  
 XX

PR 03-AUG-2000; 2000US-0632366.  
 XX

PR 21-SEP-2000; 2000US-0234687.  
 XX

PR 27-SEP-2000; 2000US-0236359.  
 XX

PR 04-OCT-2000; 2000GB-0024263.  
 XX

PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX

DR WPI; 2001-488897/53.  
 XX

PT Human genome-derived single exon nucleic acid probes useful for  
 XX analyzing gene expression in human placenta -

PS Claim 27; SEQ ID No 29540; 654bp; English.  
 XX

XX The present invention relates to single exon nucleic acid probes (SENPs;  
 CC see A113135-A1157546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders.

SQ Sequence 22 AA;

Query Match 1.8%; Score 8; DB 22; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 437 DDDDDDDG 444  
 Db 7 DDDDDDDG 14

## RESULT 50

AA104504

ID	AA04504	standard; Protein; 22 AA.
AC	AA04504;	
XX		
DT	09-OCT-2001	(first entry)
XX		
DE	Peptide #3186 encoded by probe for measuring breast gene expression.	
XX		
KW	Probe; human; breast disease; breast cancer; development disorder; inflammatory disease; proliferative breast disease; non-carcinoma tumour.	
XX		
OS	Homo sapiens.	
PN	W0200157270-A2.	
XX		
PD	09-AUG-2001.	
XX		
PF	29-JAN-2001; 2001WO-US00661.	
XX		
PR	04-FEB-2000; 2000US-0180312.	
XX	26-MAY-2000; 2000US-0207456.	
PR	30-JUN-2000; 2000US-0608408.	
XX	03-AUG-2000; 2000US-0632366.	
PR	21-SEP-2000; 2000US-0234687.	
XX	27-SEP-2000; 2000US-0236359.	
PR	04-OCT-2000; 2000GB-0024263.	
XX		
PA	(MOLE-) MOLECULAR DYNAMICS INC.	
XX		
PI	Penn SG, Hanzel DK, Chen W, Rank DR;	
XX		
DR	WPI; 2001-476286/51.	
XX		
PT	Novel single exon nucleic acid probe used to measuring gene expression	
XX	in a human breast -	
XX		
PS	Claim 27; SEQ ID No 13244; 322pp; English.	
XX		
CC	The present invention relates to novel single exon nucleic acid probes	
CC	(see AA100010-AA110067). The present sequence is a peptide encoded by one	
CC	such probe. The probes are useful for measuring human gene expression in	
CC	a human breast sample, where the probe hybridises at high stringency to a	
CC	nucleic acid expressed in the human breast. The probes are useful for	
CC	predicting, diagnosing, grading, staging, monitoring and prognosing	
CC	diseases of the human breast, particularly those diseases with polygenic	
CC	aetiology. The diseases include: breast cancer, disorders of development,	
CC	inflammatory diseases of the breast, fibrocystic changes, proliferative	
CC	breast disease and non-carcinoma tumours.	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at ftp.wipo.int/pub/published_pct_sequences.	
SQ	Sequence 22 AA;	
QY	Query Match 1.8%; Score 8; DB 22; Length 22;	
DB	Best Local Similarity 100.0%; Pred. No. 5;	
	Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0.	
	437 DDDDDDDG 444	
	7 DDDDDDDG 14	
RESULT 51		
AA062619		
ID	AA062619 standard; peptide; 22 AA.	
XX		
AC	AA062619;	
XX		
DT	23-JUL-2001 (first entry)	
XX		
DE	Soybean lunasin peptide variant (residues 22-43).	
XX		

KW	Lunasin; cancer; soybean; variant; carcinogen; viral; oncogene.
XX	
OS	Glycine max.
XX	
PN	WO200134808-A2.
XX	
PD	17-MAY-2001.
XX	
PF	12-NOV-2000; 2000WO-US31211.
XX	
PR	12-NOV-1999; 99US-0165334.
XX	
PA	(FILG-) FILGEN BIOSCIENCES INC.
XX	
F1	Galvez AF;
XX	
DR	WPI; 2001-343605/36.
XX	
PT	Large scale production of lunasin, a cancer preventive peptide from
XX	soybean, by preparing lunasin gene constructs using Pichia expression
PT	vectors, optimizing lunasin expression parameters and purifying lunasin
XX	
PT	Claim 7; Page 62; 67pp; English.
XX	
CC	The invention provides a method of producing lunasin, a cancer preventive
CC	peptide from soybean, or its active variant by recombinant DNA technology
CC	in large quantities. The method comprises preparing lunasin gene
CC	constructs or its variants using protein expression vectors, optimizing
CC	lunasin expression parameters for large-scale production and performing
CC	a series of isolation and purification to obtain large quantities of
CC	biologically active recombinant lunasin peptide. The lunasin peptide
CC	exhibits inhibitory effects against malignant transformation of cells
CC	induced by chemical carcinogens and viral oncogenes. Sequences
CC	AAB65614-24 represent soybean lunasin variants.
XX	
SQ	Sequence 22 AA;
QY	436 GDDDDDD 443
DB	13 GDDDDDD 20
RESULT 52	
ABG38546	
ID	ABG38546 standard; Peptide; 22 AA.
XX	
AC	ABG38546;
XX	
DT	19-AUG-2002 (first entry)
XX	
DE	Human peptide encoded by genome-derived single exon probe SEQ ID 28211.
XX	
KW	Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW	chronic obstructive pulmonary disease; interstitial lung disease;
KW	familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW	tubercular sclerosis; Gaucher's disease; Niemann-Pick disease;
KW	Hemansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW	pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW	pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW	primary ciliary dyskinesia; pulmonary hypertension;
KW	hyaline membrane disease.
XX	
OS	Homo sapiens.
XX	
PN	WO200186003-A2.
XX	
PD	15-NOV-2001.
XX	

PF 30-JAN-2001; 2001WO-US00665.  
 XX 04-FEB-2000; 2000US-180312P.  
 PR 26-MAY-2000; 2000US-207456P.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-063236P.  
 PR 21-SEP-2000; 2000US-2234687P.  
 PR 27-SEP-2000; 2000US-236359P.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 PI Penn SG, Hanzel DK, Chen W, Rank DR,  
 XX WPI, 2002-114183/15.  
 DR  
 PT Spatially-addressable set of single exon nucleic acid probes, used to  
 PT measure gene expression in human lung samples -  
 XX  
 Claim 27; SEQ ID No 28211; 634bp; English.

The invention relates to a spatially-addressable set of single exon nucleic acid probes for measuring gene expression in a sample derived from human lung comprising single exon nucleic acid probes having one of 12614 nucleic acid sequences mentioned in the specification, or their complements or the 12387 open reading frames derived from the 12614 probes. Also included are a microarray comprising the novel set of probes; the novel set of probes which hybridize at high stringency to a nucleic acid expressed in the human lung; measuring gene expression to a sample derived from human lung, comprising (a) contacting the array with a collection of detectably labeled nucleic acids derived from human lung mRNA, and (b) measuring the label detectably bound to each probe of the array; identifying exons in a eukaryotic genome, comprising (a) algorithmically predicting at least one exon from genomic sequences of the eukaryote; and (b) detecting specific hybridization of detectably labeled nucleic acids from eukaryote lung mRNA, to a single exon probe, having a fragment identical to the predicted exon, the probe is included in the above mentioned microarray; assigning exons to a single gene, comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridization to a single exon microarray having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary haemorrhoidosis, pulmonary histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension and hyaline membrane disease. The present sequence is a peptide/protein encoded by a single exon probe of the invention.  
 Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at  
 ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 22 AA;

Query March

Best Local Similarity 1.8%; Score 8; DB 23; Length 22;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444

DB 7 DDDDDDDG 14

RESULT 53  
 AAB83395  
 ID AAB83395 standard; peptide; 25 AA.  
 XX  
 AC AAB83395;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Lunasin related peptide #4.  
 XX  
 KW Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;  
 KW prostate; liver; kidney; therapy; soybean.  
 OS Glycine max.  
 XX  
 FN WO200172784-A2.  
 PD 04-OCT-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09453.  
 XX  
 PR 24-MAR-2000; 2000US-0534705.  
 XX  
 PA (FTIG-) FILGEN BIOSCIENCES INC.  
 XX  
 PI Galvez AF;  
 XX  
 DR WPI; 2001-648438/74.  
 XX  
 PT Treatment or prevention of cancer by administering a lunasin peptide  
 PT which binds to non-acetylated histones H3 and H4 and prevents their  
 PT acetylation in hypoacetylated repressed chromatin -  
 XX  
 PS Disclosure; Page 47; 49pp; English.  
 XX  
 CC This sequence represents a lunasin related peptide used in the method of  
 CC the invention. The method is for the treatment or prevention of cancer,  
 CC comprising administration of a lunasin peptide or its active fragment or  
 CC analog. The invention is used to treat or prevent cancer, particularly  
 CC in a human. The cancers that can be treated and prevented include  
 CC those of the colon, upper gastrointestinal tract, breast, prostate,  
 CC liver, kidney or any other internal organs or tissues.  
 XX  
 SQ Sequence 25 AA;

Query March

Best Local Similarity 1.8%; Score 8; DB 23; Length 25;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443

DB 16 GDDDDDD 23

RESULT 54

AAB29408

ID AAB29408 standard; peptide; 31 AA.

XX AAB29408;

DT 01-FEB-2002 (first entry)

XX Peptide #2059 encoded by breast cell single exon nucleic acid probe.

DE Human; microarray; single exon probe; gene expression; breast;

KW disease; cancer.

XX Homo sapiens.

OS WO200157271-A2.

PN 09-AUG-2001.

PD

```

XX PF 30-JAN-2001; 2001WO-US00662.
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX PA
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX P1 Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-496933/54.
XX PT New spatially-addressable set of single exon nucleic acid probes,
XX PR useful for measuring gene expression in sample derived from human
XX breast, comprises number of single exon nucleic acid probes -
XX Claim 27; SEQ ID NO 12376; 327bp + sequence listing; English.
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human breast and BR 474 cells. The method involves contacting
XX the probes with a collection of detectably labelled nucleic acids
XX derived from mRNA of human breast, and then measuring the label
XX bound to each probe of the microarray. The probes are useful for
XX verifying the expression of regions of genomic DNA predicted to
XX encode proteins. They are useful for gene discovery, and for
XX determining predisposition and/or prognosing the toxicity of chemical
XX expression analysis. The microarray of this invention presents a far greater
XX diversity of probes for measuring gene expression, with far less bias
XX than expressed sequence tag microarrays. The method is suitable for
XX rapid production of functional information from genomic sequence. The
XX present sequence is a peptide encoded by a single exon nucleic acid
XX probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 31 AA;
XX Query Match 1.8%; Score 8; DB 22; Length 31;
XX Best Local Similarity 100.0%; Pred No. 6.7;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX 437 DDDDDDDG 444
XX |||||
XX 16 DDDDDDDG 23
RESULT 55
ID ABB34590 standard; Peptide; 31 AA.
XX AC ABB34590;
XX DT 04-FEB-2002 (first entry)
XX PEptide #2096 encoded by human foetal liver single exon probe.
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX OS Homo sapiens.
XX PN WO200157277-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00669.
XX
```

PR	04-FEB-2000;	2000US-0180312.	
PR	26-MAY-2000;	2000US-0207456.	
PR	30-JUN-2000;	2000US-0608408.	
PR	03-AUG-2000;	2000US-0632366.	
PR	21-SEP-2000;	2000US-0234687.	
PR	27-SEP-2000;	2000US-0236359.	
PR	04-OCT-2000;	2000GB-0024263.	
XX			
PA	(MOLE-)	MOLECULAR DYNAMICS INC.	
PI	Penn SG,	Hanzel DK,	Chen W,
XX			Rank DR;
DR	WPI;	2001-483447/52.	
XX			
PT	Human genome-derived	single exon nucleic acid probes useful for	
XX	analyzing gene expression	in human fetal liver -	
PS	Claim 27;	SEQ ID NO 27225;	639pp + sequence listing; English.
XX			
CC	The invention relates	to a single exon nucleic acid probe for	
CC	measuring human gene	expression in a sample derived from human foetal	
CC	liver. The single exon	nucleic acid probes may be used for predicting,	
CC	measuring and displaying	gene expression in samples derived from human	
CC	fetal liver. The present	sequence is a peptide encoded by a single exon	
CC	nucleic acid probe of the	invention.	
CC	Note: The sequence data	for this patent did not form part of the	
CC	printed specification,	but was obtained in electronic format directly	
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences.		
XX			
SQ	Sequence	31 AA;	
QY	Query Match	1.8%;	Score 8;
	Best Local Similarity	100.0%;	Pred. No. 6.7;
	Matches	8;	Conservative
		0;	Mismatches
		0;	Indels
		0;	Gaps
DB	437 DDDDDDDG	444	
	16 DDDDDDDG	23	
RESULT 56			
ABBI19996			
ID	ABBI19996	standard;	Protein;
XX			31 AA.
AC	ABBI19996;		
XX			
DE	23-JAN-2002	(first entry)	
XX			
DE	Protein #1995	encoded by probe for	measuring heart cell gene expression.
KM	Human; gene expression;	heart; microarray; vascular system;	
KW	cardiovascular disease;	hypertension; cardiac arrhythmia;	
KW	congenital heart disease.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200157274-A2.		
XX			
PD	09-AUG-2001.		
XX			
PF	30-JAN-2001;	2001WO-US00666.	
XX			
PR	04-FEB-2000;	2000US-0180312.	
PR	26-MAY-2000;	2000US-0207456.	
PR	30-JUN-2000;	2000US-0608408.	
PR	03-AUG-2000;	2000US-0632366.	
PR	21-SEP-2000;	2000US-0234687.	
PR	27-SEP-2000;	2000US-0236359.	
PR	04-OCT-2000;	2000GB-0024263.	
XX			
PA	(MOLE-)	MOLECULAR DYNAMICS INC.	
PI	Penn SG,	Hanzel DK,	Chen W,
XX			Rank DR;
DR	WPI;	2001-483447/52.	
XX			
PT	Human genome-derived	single exon nucleic acid probes useful for	
XX	analyzing gene expression	in human fetal liver -	
PS	Claim 27;	SEQ ID NO 27225;	639pp + sequence listing; English.
XX			
CC	The invention relates	to a single exon nucleic acid probe for	
CC	measuring human gene	expression in a sample derived from human foetal	
CC	liver. The single exon	nucleic acid probes may be used for predicting,	
CC	measuring and displaying	gene expression in samples derived from human	
CC	fetal liver. The present	sequence is a peptide encoded by a single exon	
CC	nucleic acid probe of the	invention.	
CC	Note: The sequence data	for this patent did not form part of the	
CC	printed specification,	but was obtained in electronic format directly	
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences.		
XX			
SQ	Sequence	31 AA;	
QY	Query Match	1.8%;	Score 8;
	Best Local Similarity	100.0%;	Pred. No. 6.7;
	Matches	8;	Conservative
		0;	Mismatches
		0;	Indels
		0;	Gaps
DB	437 DDDDDDDG	444	
	16 DDDDDDDG	23	

XX WPI; 2001-488899/53.  
XX Single exon nucleic acid probes for analyzing gene expression in human  
PT hearts -  
XX  
PS Claim 15; SEQ ID No 21766; 530bp; English.  
XX  
CC The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart (see  
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarray.  
CC By measuring gene expression, the probes are useful for predicting,  
CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 31 AA;  
XX  
Query Match 1.8%; Score 8; DB 22; Length 31;  
Best Local Similarity 100.0%; Pred. No. 6.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 437 DDDDDDDG 444  
DB 16 DDDDDDDG 23  
XX  
RESULT 57  
AAM5379  
ID AAM5379 standard; Protein; 31 AA.  
XX  
AC AAM5379;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 27484.  
XX  
KW Human; brain expressed exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer.  
XX  
OS Homo sapiens.  
XX  
PN WO200157275-A2.  
XX  
PD 09-AUG-2001.  
XX  
PI 30-JAN-2001; 2001WO-US00667.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-483446/52.  
XX  
XX Single exon nucleic acid probes for analyzing gene expression in human  
PT brains -  
XX  
PS Example 4; SEQ ID NO: 27484; 650bp + Sequence Listing; English.  
XX

CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancer. The present sequence is a protein encoded by one of  
CC the probes of the invention.  
XX  
SQ Sequence 31 AA;  
XX  
Query Match 1.8%; Score 8; DB 22; Length 31;  
Best Local Similarity 100.0%; Pred. No. 6.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 437 DDDDDDDG 444  
DB 16 DDDDDDDG 23  
XX  
RESULT 58  
AAM6774  
ID AAM6774 standard; Protein; 31 AA.  
XX  
AC AAM6774;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 28080.  
XX  
KW Human; bone marrow expressed exon; gene expression analysis; probe;  
KW microarray; cancer; leukemia; lymphoma; myeloma.  
XX  
OS Homo sapiens.  
XX  
PN WO200157276-A2.  
XX  
PD 09-AUG-2001.  
XX  
PI 30-JAN-2001; 2001WO-US00668.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488900/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human bone marrow -  
XX  
PS Example 4; SEQ ID NO: 28080; 658bp + Sequence Listing; English.  
XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC bone marrow. They can be used to measure gene expression in bone marrow  
CC samples, which may enable the improved diagnosis and treatment of cancers  
CC such as lymphoma, leukemia and myeloma. The present sequence is a  
CC protein encoded by one of the probes of the invention.  
XX  
SQ Sequence 31 AA;  
XX  
Query Match 1.8%; Score 8; DB 22; Length 31;  
Best Local Similarity 100.0%; Pred. No. 6.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 437 DDDDDDDG 444

Db 16 DDDDDDG 23

## RESULT 59

ID AAM15583 standard; Protein; 31 AA.

AC AAM15583;

DT 12-OCT-2001 (first entry)

DE Peptide #2017 encoded by probe for measuring cervical gene expression.

KW Probe; human; microarray; gene expression; cervical epithelial cell;

KM cervical cancer.

OS Homo sapiens.

PN WO200157278-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00670.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488901/53.

PT Human genome-derived single exon nucleic acid probes useful for

analyzing gene expression in human cervical epithelial cells -

PS Claim 27; SEQ ID No 20409; 487bp; English.

CC The present invention relates to human single exon nucleic acid probes

(SENP: see AAI10068-AA128459). The present sequence is a peptide encoded

by one such probe. The SENPs are derived from human HeLa cells. The SENPs

can be used to produce a single exon microarray, which can be used for

measuring human gene expression in a sample derived from human cervical

epithelial cells. By measuring gene expression, the probes are therefore

useful in grading and/or staging of diseases of the cervix, notably

cervical cancer.

CC Note: The sequence data for this patent did not form part of the printed

specification, but was obtained in electronic format directly from WIPO

at ftp.wipo.int/pub/published\_pct\_sequences.

CC at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 31 AA;

Query Match 1.8%; Score 8; DB 22; Length 31;

Best Local Similarity 100.0%; Pred. No. 6.7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDG 444

Db 16 DDDDDDG 23

## RESULT 60

ID AAM28075 standard; Protein; 31 AA.

AC AAM28075;

DT 17-OCT-2001 (first entry)

DE Peptide #2112 encoded by probe for measuring placental gene expression.

KW Probe; microarray; human; placenta; antenatal diagnosis;

KM genetic disorder.

OS Homo sapiens.

PN WO200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00663.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-48897/53.

PT Human genome-derived single exon nucleic acid probes useful for

analyzing gene expression in human placenta -

PS Claim 27; SEQ ID No 28344; 654bp; English.

CC The present invention relates to single exon nucleic acid probes (SENP:

see AAI1315-AA157546). The present sequence is a peptide encoded by one

such probe. The probes are useful for producing a microarray for

predicting, measuring and displaying gene expression in samples derived

from human placenta. The probes are useful for antenatal diagnosis of

human genetic disorders.

CC human genetic disorders.

CC human genetic disorders.

CC human genetic disorders.

CC human genetic disorders.

CC human genetic disorders.

CC human genetic disorders.

CC human genetic disorders.

CC human genetic disorders.

CC human genetic disorders.

CC human genetic disorders.

CC human genetic disorders.

CC human genetic disorders.

CC human genetic disorders.



PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-476286/51.  
XX  
XX Novel single exon nucleic acid probe used to measuring gene expression  
XX in a human breast -  
XX  
XX Claim 27; SEQ ID No 12067; 322pp; English.  
XX  
XX The present invention relates to novel single exon nucleic acid probes  
XX (see A100010-A110067). The present sequence is a peptide encoded by one  
XX such probe. The probes are useful for measuring human gene expression in  
XX a human breast sample, where the probe hybridizes at high stringency to a  
XX nucleic acid expressed in the human breast. The probes are useful for  
XX predicting, diagnosing, grading, staging, monitoring and prognosis  
XX diseases of the human breast, particularly those diseases with polygenic  
XX etiology. The diseases include: breast cancer, disorders of development,  
XX inflammatory diseases of the breast, fibrocystic changes, proliferative  
XX breast disease and non-carcinoma tumours.  
XX Note: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 31 AA;  
XX  
XX Query Match 1.8%; Score 8; DB 22; Length 31;  
XX Best Local Similarity 100.0%; Pred. No. 6.7;  
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 437 DDDDDDDG 444  
XX 16 DDDDDDDG 23  
XX  
XX RESULT 62  
XX ABB37322  
XX ABB37322 standard; Peptide; 31 AA.  
XX  
XX ABB37322;  
XX  
XX 19-AUG-2002 (first entry)  
XX  
XX Human peptide encoded by genome-derived single exon probe SEQ ID 26987.  
XX  
XX Human; single exon probe; asthma; lung cancer; COPD; ILD;  
XX chronic obstructive pulmonary disease; interstitial lung disease;  
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;  
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
XX Heremansky-Pudlak syndrome; sarcoidosis; pulmonary haemostiderosis;  
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;  
XX primary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
XX primary ciliary dyskinesia; pulmonary hypertension;  
XX hyaline membrane disease.  
XX  
XX Homo sapiens.  
XX  
XX WO200186003-A2.  
XX  
XX 15-NOV-2001.  
XX  
XX 30-JAN-2001; 2001WO-US00665.  
XX  
XX 04-FEB-2000; 2000US-180312P.  
XX

PR 26-MAY-2000; 2000US-207456P.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-234687P.  
PR 27-SEP-2000; 2000US-236359P.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2002-114183/15.  
XX  
XX Spatially-addressable set of single exon nucleic acid probes, used to  
XX measure gene expression in human lung samples -  
XX  
XX Claim 27; SEQ ID No 26987; 634pp; English.  
XX  
XX The invention relates to a spatially-addressable set of single exon  
XX nucleic acid probes for measuring gene expression in a sample derived  
XX from human lung comprising single exon nucleic acid probes having one of  
XX 12614 nucleic acid sequences mentioned in the specification, or their  
XX complements or the 12387 open reading frames derived from the 12614  
XX probes. Also included are a microarray comprising the novel set of  
XX probes; the novel set of probes which hybridize at high stringency to a  
XX nucleic acid expressed in the human lung; measuring gene expression in a  
XX sample derived from human lung, comprising (a) contacting the array with  
XX a collection of detectably labeled nucleic acid derived from human lung  
XX mRNA, and (b) measuring the label detectably bound to each probe of  
XX the array; identifying exons in a eukaryotic genome, comprising  
XX (a) algorithmically predicting at least one exon from genomic sequences  
XX of the eukaryote; and (b) detecting specific hybridisation of detectably  
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
XX in the above mentioned microarray; assigning exons to a single gene,  
XX comprising (a) identifying exons from genomic sequence by the method  
XX above and (b) measuring the expression of each of the exons in several  
XX tissues and/or cell types using hybridisation to a single exon  
XX microarrays having a probe with the exon, where a common pattern of  
XX expression of the exons in the tissues and/or cell types indicates that  
XX of the exons should be assigned to a single gene; a peptide comprising one  
XX of 12011 sequences, mentioned in the specification, or encoded by the  
XX probes/open reading frames (ORF). The probes are used for gene  
XX expression analysis, and for identifying exons in a gene, particularly  
XX using human lung derived mRNA and for the study of lung diseases  
XX such as asthma, lung cancer, chronic obstructive pulmonary disease  
XX (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
XX fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
XX Niemann-Pick disease, Heremansky-Pudlak syndrome, sarcoidosis, pulmonary  
XX haemostiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
XX pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic  
XX pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension  
XX and hyaline membrane disease. The present sequence is a peptide/protein  
XX encoded by a single exon probe of the invention.  
XX Note: The sequence data for this patent did not form part  
XX of the printed specification, but was obtained in electronic  
XX format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 31 AA;  
XX  
XX Query Match 1.8%; Score 8; DB 23; Length 31;  
XX Best Local Similarity 100.0%; Pred. No. 6.7;  
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 437 DDDDDDDG 444  
XX 16 DDDDDDDG 23  
XX  
XX RESULT 63  
XX AAB83381  
XX AAB83381 standard; peptide; 41 AA.  
XX

XX AAB83381;  
AC 26-MAR-2002 (first entry)  
XX  
DT Lunasin fragment #2.  
XX  
DE Lunasin fragment #2.  
XX  
XX Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;  
XX prostate; liver; kidney; therapy; soybean.  
XX  
XX Glycine max.  
XX OS  
XX WO200172784-A2.  
XX  
XX 04-OCT-2001.  
XX  
XX 23-MAR-2001; 2001WO-US09453.  
XX  
XX 24-MAR-2000; 2000US-0534705.  
XX  
XX (FIG-) FILGEN BIOSCIENCES INC.  
XX Galvez AF;  
XX  
XX WPI, 2001-648438/74.  
XX  
XX Treatment or prevention of cancer by administering a lunasin peptide  
XX which binds to non-acetylated histones H3 and H4 and prevents their  
XX acetylation in hypoacetylated repressed chromatin  
XX  
XX Claim 7; Page 41; 49pp; English.  
XX  
XX This sequence represents a lunasin peptide used in the method of the  
XX invention. The method is for the treatment or prevention of cancer,  
XX comprising administration of a lunasin peptide or its active fragment or  
XX analog. The invention is used to treat or prevent cancer, particularly  
XX in a human. The cancers that can be treated and prevented include  
XX those of the colon, upper gastrointestinal tract, breast, prostate,  
XX liver, kidney or any other internal organs or tissues.  
XX  
SQ Sequence 41 AA;  
Query Match 1.8%; Score 8; DB 22; Length 41;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 436 GDDDDDD 443  
34 GDDDDDD 41  
RESULT 64  
AAB62615  
ID AAB62615 standard; peptide; 41 AA.  
XX  
AC AAB62615;  
XX  
DT 23-JUL-2001 (first entry)  
XX  
XX Soybean lunasin peptide variant (residues 1-41).  
XX  
XX Lunasin; cancer; soybean; variant; carcinogen; viral; oncogene.  
XX  
XX Glycine max.  
XX OS  
XX WO200134808-A2.  
XX  
XX 17-MAY-2001.  
XX  
XX 12-NOV-2000; 2000WO-US31211.  
XX  
XX 12-NOV-1999; 99US-0165334.  
XX

PA (FIG-) FILGEN BIOSCIENCES INC.  
XX  
XX Galvez AF;  
XX  
XX WPI, 2001-343605/36.  
XX  
XX Large scale production of lunasin, a cancer preventive peptide from  
XX soybean, by preparing lunasin gene constructs using Pichia expression  
XX vectors, optimizing lunasin expression parameters and purifying lunasin  
XX  
XX Claim 7; Page 60; 67pp; English.  
XX  
XX The invention provides a method of producing lunasin, a cancer preventive  
XX peptide from soybean, or its active variant by recombinant DNA technology  
XX in large quantities. The method comprises preparing lunasin gene  
XX constructs or its variants using protein expression vectors, optimizing  
XX lunasin expression parameters for large-scale production and performing  
XX a series of isolation and purification to obtain large quantities of  
XX biologically active recombinant lunasin peptide. The lunasin peptide  
XX exhibits inhibitory effects against malignant transformation of cells  
XX induced by chemical carcinogens and viral oncogenes. Sequences  
XX AAB62614-24 represent soybean lunasin variants.  
XX  
SQ Sequence 41 AA;  
Query Match 1.8%; Score 8; DB 22; Length 41;  
Best Local Similarity 100.0%; Pred. No. 8.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 436 GDDDDDD 443  
34 GDDDDDD 41  
RESULT 65  
AAB83380  
ID AAB83380 standard; peptide; 42 AA.  
XX  
XX AAB83380;  
XX AC  
XX 26-MAR-2002 (first entry)  
XX  
XX Lunasin fragment #1.  
XX  
XX Lunasin; cancer; human; colon; upper gastrointestinal tract; breast;  
XX prostate; liver; kidney; therapy; soybean.  
XX  
XX Glycine max.  
XX OS  
XX WO200172784-A2.  
XX  
XX 04-OCT-2001.  
XX  
XX 23-MAR-2001; 2001WO-US09453.  
XX  
XX 24-MAR-2000; 2000US-0534705.  
XX  
XX (FIG-) FILGEN BIOSCIENCES INC.  
XX  
XX Galvez AF;  
XX  
XX WPI, 2001-648438/74.  
XX  
XX Treatment or prevention of cancer by administering a lunasin peptide  
XX which binds to non-acetylated histones H3 and H4 and prevents their  
XX acetylation in hypoacetylated repressed chromatin  
XX  
XX Claim 7; Page 40; 49pp; English.  
XX  
XX This sequence represents a lunasin peptide used in the method of the  
XX invention. The method is for the treatment or prevention of cancer,  
XX comprising administration of a lunasin peptide or its active fragment or

CC analog. The invention is used to treat or prevent cancer, particularly  
CC in a human. The cancers that can be treated and prevented include  
CC those of the colon, upper gastrointestinal tract, breast, prostate,  
CC liver, kidney or any other internal organs or tissues.  
XX  
SQ Sequence 42 AA;

Query Match 1.8%; Score 8; DB 22; Length 42;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 436 GDDDDDD 443  
Db 34 GDDDDDD 41

RESULT 66  
ABB31264  
ABB31264 standard; Peptide; 42 AA.  
ABB31264;

01-FEB-2002 (first entry)

Peptide #3915 encoded by breast cell single exon nucleic acid probe.

Human; microarray; single exon probe; gene expression; breast;  
disease; cancer.

Homo sapiens.

WO200157271-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US00662.

04-FEB-2000; 2000US-0180312.

26-MAY-2000; 2000US-0207456.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0633366.

21-SEP-2000; 2000US-0234687.

27-SEP-2000; 2000US-0236359.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-496933/54.

New spatially-addressable set of single exon nucleic acid probes,  
useful for measuring gene expression in sample derived from human  
breast, comprises number of single exon nucleic acid probes -

Claim 27; SEQ ID NO 14232; 327bp + sequence listing; English.

The invention relates to a spatially-addressable set of single exon  
nucleic acid probes for measuring gene expression in a sample derived  
from human breast and BT 474 cells. The method involves contacting  
the probes with a collection of detectably labeled nucleic acids  
derived from mRNA of human breast, and then measuring the label  
bound to each probe of the microarray. The probes are useful for  
verifying the expression of regions of genomic DNA predicted to  
encode proteins. They are useful for gene discovery, and for  
determining predisposition and/or prognosing breast disease. Gene  
expression analysis is useful for assessing the toxicity of chemical  
agents on cells. The microarray of this invention presents a far greater  
diversity of probes for measuring gene expression, with far less bias  
than expressed sequence tag microarrays. The method is suitable for  
rapid production of functional information from genomic sequence. The  
present sequence is a peptide encoded by a single exon nucleic acid  
probe of the invention.

CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 42 AA;

Query Match 1.8%; Score 8; DB 22; Length 42;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 437 DDDDDDDG 444  
Db 14 DDDDDDDG 21

RESULT 67  
ABB36472  
ABB36472 standard; Peptide; 42 AA.  
ABB36472;

04-FEB-2002 (first entry)

Peptide #3978 encoded by human foetal liver single exon probe.

Human; foetal liver; gene expression; single exon nucleic acid probe.

Homo sapiens.

WO200157277-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US00669.

04-FEB-2000; 2000US-0180312.

26-MAY-2000; 2000US-0207456.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0633366.

21-SEP-2000; 2000US-0234687.

27-SEP-2000; 2000US-0236359.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-483447/52.

Human genome-derived single exon nucleic acid probes useful for  
analyzing gene expression in human fetal liver -

Claim 27; SEQ ID NO 29107; 639bp + sequence listing; English.

The invention relates to a single exon nucleic acid probe for  
measuring human gene expression in a sample derived from human foetal  
liver. The single exon nucleic acid probes may be used for predicting,  
measuring and displaying gene expression in samples derived from human  
fetal liver. The present sequence is a peptide encoded by a single exon  
nucleic acid probe of the invention.

Note: The sequence data for this patent did not form part of the  
printed specification, but was obtained in electronic format directly  
from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 42 AA;

Query Match 1.8%; Score 8; DB 22; Length 42;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDDG 444

Db 14 DDDDDDDG 21

RESULT 68  
 ABB21815  
 ID ABB21815 standard; Protein: 42 AA.  
 XX  
 AC ABB21815;  
 XX  
 DT 23-JAN-2002 (first entry)  
 XX  
 DE Protein #3814 encoded by probe for measuring heart cell gene expression.  
 XX  
 KW Human; gene expression; heart; microarray; vascular system;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157274-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PR 30-JAN-2001; 2001WO-US00666.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488899/53.  
 XX  
 PT Single exon nucleic acid probes for analyzing gene expression in human  
 PT hearts -  
 XX  
 PS Claim 15; SEQ ID No 23585; 530bp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart (see  
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
 CC probe. The probes may be used for predicting, measuring and displaying  
 CC gene expression in samples derived from the human heart via microarrays.  
 CC By measuring gene expression, the probes are useful for predicting,  
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
 CC human heart and vascular system e.g. cardiovascular disease,  
 CC hypertension, cardiac arrhythmias and congenital heart disease.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 42 AA;  
 XX  
 QY Query Match 1.8%; Score 8; DB 22; Length 42;  
 QY Best Local Similarity 100.0%; Pred. No. 8.7;  
 QY Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 437 DDDDDDDG 444  
 QY |||||  
 QY 14 DDDDDDDG 21  
 DB

RESULT 69  
 AAM57237  
 ID AAM57237 standard; Protein: 42 AA.  
 XX  
 AC AAM57237;  
 XX  
 DT 05-NOV-2001 (first entry)  
 XX  
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 29342.  
 XX  
 KW Human; brain expressed exon; gene expression analysis; probe;  
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
 KW epilepsy; cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157275-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PR 30-JAN-2001; 2001WO-US00667.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-483446/52.  
 XX  
 PT Single exon nucleic acid probes for analyzing gene expression in human  
 PT brains -  
 XX  
 PS Example 4; SEQ ID NO: 29342; 650bp + Sequence listing; English.  
 XX  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,  
 CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is a protein encoded by one of  
 CC the probes of the invention.  
 XX  
 SQ Sequence 42 AA;  
 XX  
 QY Query Match 1.8%; Score 8; DB 22; Length 42;  
 QY Best Local Similarity 100.0%; Pred. No. 8.7;  
 QY Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 437 DDDDDDDG 444  
 QY |||||  
 QY 14 DDDDDDDG 21  
 DB

RESULT 70  
 AAM69636  
 ID AAM69636 standard; Protein: 42 AA.  
 XX  
 AC AAM69636;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 29942.  
 XX  
 KW Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukemia; lymphoma; myeloma.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157276-A2.  
 XX  
 PD 09-AUG-2001.  
 XX

PF 30-JAN-2001; 2001WO-US00668.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488900/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 XX analyzing gene expression in human bone marrow -  
 XX  
 CC Example 4; SEQ ID NO: 29942; 658bp + Sequence Listing; English.  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention.  
 CC  
 SQ Sequence 42 AA;  
 XX  
 Query Match 1.8%; Score 8; DB 22; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 8.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 437 DDDDDDDG 444  
 DB 14 DDDDDDDG 21  
 XX  
 RESULT 71  
 AAM17452  
 ID AAM17452 standard; Protein; 42 AA.  
 XX  
 AC AAM17452;  
 XX  
 DT 12-OCT-2001 (first entry)  
 XX  
 PE Peptide #3886 encoded by probe for measuring cervical gene expression.  
 KM Probe: human; microarray; gene expression; cervical epithelial cell;  
 KM cervical cancer.  
 KM  
 OS Homo sapiens.  
 XX  
 PN WO200157278-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00670.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488901/53.

XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human cervical epithelial cells -  
 XX  
 PS Claim 27; SEQ ID NO 22278; 487bp; English.  
 XX  
 CC The present invention relates to human single exon nucleic acid probes  
 CC (SENP; see A110068-A1128459). The present sequence is a peptide encoded  
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pcr\_sequences.  
 XX  
 SQ Sequence 42 AA;  
 XX  
 Query Match 1.8%; Score 8; DB 22; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 8.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 437 DDDDDDDG 444  
 DB 14 DDDDDDDG 21  
 XX  
 RESULT 72  
 AAM2972  
 ID AAM2972 standard; Protein; 42 AA.  
 XX  
 AC AAM2972;  
 XX  
 DT 17-OCT-2001 (first entry)  
 XX  
 DE Peptide #4009 encoded by probe for measuring placental gene expression.  
 XX  
 KM Probe: microarray; human; placenta; antenatal diagnosis;  
 KM genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157272-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00663.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-48897/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human placenta -  
 XX  
 PS Claim 27; SEQ ID NO 30241; 654bp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes (SENP;  
 CC see A113135-A1157546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived

CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders.

XX  
SQ Sequence 42 AA;

Query Match 1.8%; Score 8; DB 22; Length 42;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
|||  
Db 14 DDDDDDDG 21

## RESULT 73

AAM05129  
ID AAM05129 standard; Protein; 42 AA.

XX  
AC AAM05129;

XX 09-OCT-2001 (first entry)

XX Peptide #3811 encoded by probe for measuring breast gene expression.

XX Probe: human; breast disease; breast cancer; development disorder;  
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.

XX Homo sapiens.

XX MO200157270-A2.

XX 09-AUG-2001.

XX 29-JAN-2001; 2001WO-US00061.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-476286/51.

XX Novel single exon nucleic acid probe used to measuring gene expression  
in a human breast -

XX Claim 27; SEQ ID No 13869; 322pp; English.

XX The present invention relates to novel single exon nucleic acid probes  
(see A1100010-A110067). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for measuring human gene expression in  
CC a human breast sample, where the probe hybridizes at high stringency to a  
CC nucleic acid expressed in the human breast. The probes are useful for  
CC predicting, diagnosing, grading, staging, monitoring and prognosis  
CC diseases of the human breast, particularly those diseases with polygenic  
CC aetiology. The diseases include: breast cancer, disorders of development,  
CC inflammatory diseases of the breast, fibrocystic changes, proliferative  
CC breast disease and non-carcinoma tumours.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WFO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 42 AA;

Query Match 1.8%; Score 8; DB 22; Length 42;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
|||  
Db 14 DDDDDDDG 21

## RESULT 74

AAB62614  
ID AAB62614 standard; peptide; 42 AA.

XX  
AC AAB62614;

XX 23-JUL-2001 (first entry)

XX Soybean lunasin peptide variant (residues 1-42).

XX Lunasin; cancer; soybean; variant; carcinogen; viral; oncogene.

XX Glycine max.

XX MO200134808-A2.

XX 17-MAY-2001.

XX 12-NOV-2000; 2000WO-US31211.

XX 12-NOV-1999; 99US-0165334.

XX (FILG-) FILGEN BIOSCIENCES INC.

XX Galvez AF;

XX WPI; 2001-343605/36.

XX Large scale production of lunasin, a cancer preventive peptide from  
PT soybean, by preparing lunasin gene constructs using Pichia expression  
PT vectors, optimizing lunasin expression parameters and purifying lunasin  
XX Claim 7; Page 60; 67pp; English.

XX The invention provides a method of producing lunasin, a cancer preventive  
XX peptide from soybean, or its active variant by recombinant DNA technology  
XX in large quantities. The method comprises preparing lunasin gene  
XX constructs or its variants using protein expression vectors, optimizing  
XX lunasin expression parameters for large-scale production and performing  
XX a series of isolation and purification to obtain large quantities of  
XX biologically active recombinant lunasin peptide. The lunasin peptide  
XX exhibits inhibitory effects against malignant transformation of cells  
XX induced by chemical carcinogens and viral oncogenes. Sequences  
XX AAB62614-24 represent soybean lunasin variants.

XX Sequence 42 AA;

Query Match 1.8%; Score 8; DB 22; Length 42;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDG 443  
|||  
Db 34 GDDDDDDG 41

## RESULT 75

ABG39258  
ID ABG39258 standard; Peptide; 42 AA.

XX  
AC ABG39258;

XX 19-AUG-2002 (first entry)

XX Human peptide encoded by genome-derived single exon probe SEQ ID 28923.

Human; single exon probe; asthma; lung cancer; COPD; ILD; chronic obstructive pulmonary disease; interstitial lung disease; familial idiopathic pulmonary fibrosis; neurofibromatosis; tuberous sclerosis; Gaucher's disease; Niemann-Pick disease; Hermansky-Pudlak syndrome; sarcoidosis; pulmonary hemosiderosis; pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome; pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia; primary ciliary dyskinesia; pulmonary hypertension; hyaline membrane disease.

Homo sapiens.

WO200186003-A2.

15-NOV-2001.

30-JAN-2001; 2001WO-US00665.

04-FEB-2000; 2000US-180312P.

26-MAY-2000; 2000US-207456P.

30-JUN-2000; 2000US-060840B.

03-AUG-2000; 2000US-063236G.

21-SEP-2000; 2000US-234687P.

27-SEP-2000; 2000US-236359P.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2002-114183/15.

Spatially-addressable set of single exon nucleic acid probes, used to measure gene expression in human lung samples -

Claim 27; SEQ ID No 28923; 634pp; English.

The invention relates to a spatially-addressable set of single exon nucleic acid probes for measuring gene expression in a sample derived from human lung comprising single exon nucleic acid probes having one of 12614 nucleic acid sequences mentioned in the specification, or their complements or the 12387 open reading frames derived from the 12614 probes. Also included are a microarray comprising the novel set of probes; the novel set of probes which hybridize at high stringency to a nucleic acid expressed in the human lung; measuring gene expression in a sample derived from human lung, comprising (a) contacting the array with a collection of detectably labeled nucleic acids derived from human lung mRNA, and (b) measuring the label detectably bound to each probe of the array; identifying exons in a eukaryotic genome, comprising (a) algorithmically predicting at least one exon from genomic sequences of the eukaryote; and (b) detecting specific hybridization of detectably labeled nucleic acids from eukaryote lung mRNA, to a single exon probe, having a fragment identical to the predicted exon, the probe is included in the above mentioned microarray; assigning exons to a single gene, comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridization to a single exon microarray having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary hemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension and hyaline membrane disease. The present sequence is a peptide/protein encoded by a single exon probe of the invention.

CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SQ Sequence 42 AA;

Query Match 1.8%; Score 8; DB 23; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 8.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
 DB 14 DDDDDDDG 21

Search completed: May 1, 2003, 20:53:02  
 Job time : 49 secs





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using bw model

Run on: May 1, 2003, 20:55:38 ; Search time 48 Seconds

(without alignments)  
817.926 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 455

Sequence: 1 MSISGASFWQIKFDDLPFE.....GDDDDDDGEEEDNDMNSE 455

Scoring table: OLIGO Gapop 60.0, Gapext 60.0

Searched: 328255 seqs, 8628665 residues

Seed size: \*

Total number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCTI\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTIS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	455	100.0	455	10	US-09-757-982-5
2	80	17.6	92	9	US-09-764-868-799
3	9	2.0	502	10	US-09-903-068-12
4	8	1.8	21	9	US-10-252-256-3
5	8	1.8	22	10	US-09-864-761-36493
6	8	1.8	25	9	US-10-252-256-5
7	8	1.8	31	10	US-09-864-761-35294
8	8	1.8	42	10	US-09-864-761-37113
9	8	1.8	43	9	US-10-252-256-1
10	8	1.8	49	9	US-09-836-392-34
11	8	1.8	236	10	US-09-809-545A-6
12	8	1.8	286	10	US-09-801-368-188
13	8	1.8	328	10	US-09-862-027-18
14	8	1.8	416	9	US-08-291-417-5
15	8	1.8	416	10	US-09-862-027-14
16	8	1.8	487	10	US-09-810-808-8
17	8	1.8	579	9	US-10-158-895-4
18	8	1.8	590	9	US-10-158-895-15
19	8	1.8	847	9	US-10-143-133-2

20	1.5	8	9	US-09-882-291-22	Sequence 22, Appl
21	1.5	9	9	US-09-882-291-29	Sequence 29, Appl
22	1.5	12	9	US-09-876-994A-173	Sequence 173, Appl
23	1.5	14	9	US-09-911-569-80	Sequence 80, Appl
24	1.5	20	10	US-09-864-761-38908	Sequence 38908, A
25	1.5	23	10	US-09-864-761-36028	Sequence 36028, A
26	1.5	24	10	US-09-864-761-34620	Sequence 34620, A
27	1.5	24	10	US-09-864-761-36142	Sequence 36142, A
28	1.5	26	10	US-09-864-761-35447	Sequence 35447, A
29	1.5	27	10	US-09-864-761-39392	Sequence 39392, A
30	1.5	30	9	US-09-911-569-87	Sequence 87, Appl
31	1.5	31	10	US-09-864-761-33555	Sequence 33555, A
32	1.5	34	10	US-09-864-761-44046	Sequence 44046, A
33	1.5	36	10	US-09-864-761-35933	Sequence 35933, A
34	1.5	37	10	US-09-864-761-43116	Sequence 43116, A
35	1.5	40	9	US-10-252-256-2	Sequence 2, Appl
36	1.5	40	10	US-09-864-761-39102	Sequence 39102, A
37	1.5	42	10	US-09-864-761-35018	Sequence 35018, A
38	1.5	42	10	US-09-864-761-35629	Sequence 35629, A
39	1.5	51	10	US-09-864-761-39869	Sequence 39869, A
40	1.5	61	10	US-09-864-761-39079	Sequence 39079, A
41	1.5	86	12	US-10-028-247-2	Sequence 2, Appl
42	1.5	91	10	US-09-873-880-4	Sequence 873, Appl
43	1.5	102	10	US-09-864-761-44209	Sequence 44209, A
44	1.5	110	10	US-10-101-464A-62215	Sequence 42215, A
45	1.5	154	9	US-10-101-464A-662	Sequence 662, Appl
46	1.5	154	10	US-09-864-761-43542	Sequence 43542, A
47	1.5	161	10	US-10-101-464A-681	Sequence 681, Appl
48	1.5	163	9	US-09-764-868-816	Sequence 816, Appl
49	1.5	170	9	US-09-764-868-845	Sequence 845, Appl
50	1.5	170	9	US-09-764-891-4540	Sequence 4540, Appl
51	1.5	174	9	US-09-764-868-833	Sequence 833, Appl
52	1.5	177	9	US-09-764-891-5028	Sequence 5028, Appl
53	1.5	177	9	US-10-101-464A-540	Sequence 540, Appl
54	1.5	177	9	US-10-101-464A-702	Sequence 702, Appl
55	1.5	207	9	US-09-738-626-4510	Sequence 4510, Appl
56	1.5	238	9	US-10-101-464A-1424	Sequence 1424, Appl
57	1.5	244	10	US-09-925-301-1424	Sequence 1063, Appl
58	1.5	250	9	US-09-764-868-1063	Sequence 77, Appl
59	1.5	250	9	US-09-955-899-77	Sequence 32, Appl
60	1.5	256	9	US-09-898-837A-332	Sequence 14, Appl
61	1.5	257	9	US-09-949-842-14	Sequence 821, Appl
62	1.5	255	9	US-09-764-868-821	Sequence 6, Appl
63	1.5	271	10	US-09-840-704-6	Sequence 27, Appl
64	1.5	275	9	US-10-274-409-5	Sequence 6, Appl
65	1.5	277	9	US-10-274-409-6	Sequence 14, Appl
66	1.5	277	9	US-10-274-409-7	Sequence 14, Appl
67	1.5	280	8	US-08-910-386A-14	Sequence 106, Appl
68	1.5	284	10	US-09-771-161A-106	Sequence 22, Appl
69	1.5	295	9	US-10-060-019-22	Sequence 5, Appl
70	1.5	303	9	US-09-991-211-5	Sequence 27, Appl
71	1.5	304	10	US-09-866-582-27	Sequence 61, Appl
72	1.5	324	10	US-09-746-801A-61	Sequence 5783, A
73	1.5	325	10	US-09-815-242-5783	Sequence 12693, A
74	1.5	335	10	US-09-815-242-13693	Sequence 12779, A
75	1.5	335	10	US-09-815-242-13779	Sequence 13070, A
76	1.5	335	10	US-09-815-242-11070	Sequence 1268, Appl
77	1.5	329	10	US-09-925-300-1268	Sequence 116, Appl
78	1.5	333	10	US-09-801-366-116	Sequence 6, Appl
79	1.5	337	9	US-10-026-021-6	Sequence 2, Appl
80	1.5	337	9	US-09-944-413-2	Sequence 2, Appl
81	1.5	339	9	US-09-944-403-2	Sequence 2, Appl
82	1.5	339	9	US-09-944-896-2	Sequence 2, Appl
83	1.5	339	9	US-09-944-944-2	Sequence 2, Appl
84	1.5	339	9	US-09-944-907-2	Sequence 2, Appl
85	1.5	339	9	US-10-028-072-328	Sequence 328, Appl
86	1.5	339	9	US-10-121-094-328	Sequence 328, Appl
87	1.5	339	9	US-10-123-044-328	Sequence 328, Appl
88	1.5	379	9	US-10-140-470-328	Sequence 328, Appl
89	1.5	379	9	US-10-175-746-328	Sequence 328, Appl
90	1.5	379	9	US-10-175-746-328	Sequence 328, Appl
91	1.5	379	9	US-10-175-746-328	Sequence 328, Appl
92	1.5	379	9	US-10-175-746-328	Sequence 328, Appl

93 7 1.5 379 9 US-10-176-921-328 Sequence 328, App  
94 7 1.5 379 9 US-10-137-865-328 Sequence 328, App  
95 7 1.5 379 9 US-10-140-474-328 Sequence 328, App  
96 7 1.5 379 9 US-10-142-431-328 Sequence 328, App  
97 7 1.5 379 9 US-10-143-114-328 Sequence 328, App  
98 7 1.5 379 9 US-10-140-002-328 Sequence 328, App  
99 7 1.5 379 9 US-10-142-419-328 Sequence 328, App  
100 7 1.5 379 9 US-10-123-262-328 Sequence 328, App

## ALIGNMENTS

RESULT 1  
US-09-757-982-5  
; Sequence 5, Application US/09757982  
; Patent No. US20020094559A1  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: NMI-050  
; CURRENT APPLICATION NUMBER: US/09/757,982  
; PRIOR FILING DATE: 2001-01-10  
; PRIOR APPLICATION NUMBER: 09/163,115  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-757-982-5

Query Match 100.0%; Score 455; DB 10; Length 455;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSLCASFVQIKFDDLOFFENCGGSGFSGVYAKMISQDKYAVVKKLKIKEAEILSVL 60  
DB 1 MSSLCASFVQIKFDDLOFFENCGGSGFSGVYAKMISQDKYAVVKKLKIKEAEILSVL 60  
QY 61 SHRNIIOFGVILPEPPNGIVTEYASIGSLYDIYINSNRSEEMDMHIMTATDVAKGMY 120  
DB 61 SHRNIIOFGVILPEPPNGIVTEYASIGSLYDIYINSNRSEEMDMHIMTATDVAKGMY 120  
QY 121 LMEAPVAVIHRDLKSRNVVIAADGVLCDFGASRFNNHTTHMSIVGTFPMMADEVIOS 180  
DB 121 LMEAPVAVIHRDLKSRNVVIAADGVLCDFGASRFNNHTTHMSIVGTFPMMADEVIOS 180  
QY 181 LPVSETCDYISGVVLMWLTREVPKGLGQVAMLVVEKNERLTIPSCRSFAEILH 240  
DB 181 LPVSETCDYISGVVLMWLTREVPKGLGQVAMLVVEKNERLTIPSCRSFAEILH 240  
QY 241 OCMEADAKRPSFKOITISLESMSNDTSLPDKNSFLHNKAEWRCIEATLERLKLKLEND 300  
DB 241 OCMEADAKRPSFKOITISLESMSNDTSLPDKNSFLHNKAEWRCIEATLERLKLKLEND 300  
QY 301 ISFKQELKERERRRLKMEQKLTESQNTPLLLPLAARMEESYFESKTESNSAEMSCOI 360  
DB 301 ISFKQELKERERRRLKMEQKLTESQNTPLLLPLAARMEESYFESKTESNSAEMSCOI 360  
QY 361 TATSNEEGGQNPSTQAMLMGFGDI FSNMKA GAVHSGQIMQAKONSSKTTSKRQK 420  
DB 361 TATSNEEGGQNPSTQAMLMGFGDI FSNMKA GAVHSGQIMQAKONSSKTTSKRQK 420  
QY 421 KYNMALGFSDPDLSEGDGDDDDGEEEDNDMDNSE 455  
DB 421 KYNMALGFSDPDLSEGDGDDDDGEEEDNDMDNSE 455

RESULT 2  
US-09-764-868-799  
; Sequence 799, Application US/09764868

; Patent No. US20020168711A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT232  
; CURRENT APPLICATION NUMBER: US/09/764,868  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1510  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 799  
; LENGTH: 92  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (4)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-868-799

Query Match 17.6%; Score 80; DB 9; Length 92;  
Best Local Similarity 100.0%; Pred. No. 3,4e-66;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 AEILSVLSHRNIIOFGVILPEPPNGIVTEYASIGSLYDIYINSNRSEEMDMHIMTATD 113  
DB 13 AEILSVLSHRNIIOFGVILPEPPNGIVTEYASIGSLYDIYINSNRSEEMDMHIMTATD 72  
QY 114 VAKGMHYLMEAPVAVIHRD 133  
DB 73 VAKGMHYLMEAPVAVIHRD 92

RESULT 3  
US-09-903-068-12  
; Sequence 12, Application US/09903068  
; Patent No. US20020123139A1  
; GENERAL INFORMATION:

APPLICANT: Miyazono, Kohel; Dijke, Peter Ten;

TITLE OF INVENTION: Activin Receptor-Like Kinase, Helidin, Carl-Henrik

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage

COMPUTER: IBM

OPERATING SYSTEM: PC-DOS

SOFTWARE: WordPerfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/903,068

FILING DATE: 11-Jul-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/679,187

FILING DATE: <Unknown>

APPLICATION NUMBER: PCT/GB93/02367

FILING DATE: 17-No. US20020123139A1member-1993

APPLICATION NUMBER: 9224057.1

FILING DATE: 17-No. US20020123139A1member-1992

APPLICATION NUMBER: 9304677.9

FILING DATE: 8-March-1993

APPLICATION NUMBER: 9304680.3

FILING DATE: 8-March-1993

APPLICATION NUMBER: 9311047.6

FILING DATE: 28-May-1993

APPLICATION NUMBER: 9313763.6

FILING DATE: 2-July-1993

APPLICATION NUMBER: 9136099.2

FILED DATE: 3-August-1993  
APPLICATION NUMBER: 9321344.5  
FILING DATE: 15-October-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Kohlei, Vinet  
REGISTRATION NUMBER: 37,003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3864  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 502 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-09-903-068-12

Query Match  
Best Local Similarity 2.0%; Score 9; DB 10; Length 502;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 HRDLKSRNV 139  
Db 327 HRDLKSRNV 335

RESULT 4  
US-10-252-256-3  
Sequence 3, Application US/10252256  
Publication No. US2003002765A1  
GENERAL INFORMATION:  
APPLICANT: GALVEZ, ALFREDO F.  
TITLE OF INVENTION: THERAPEUTIC PEPTIDES HAVING A MOTIF THAT BINDS  
SPECIFICALLY TO NON-ACETYLATED H3 AND H4 HISTONES FOR  
CANCER THERAPY  
FILE REFERENCE: 3729.02  
CURRENT APPLICATION NUMBER: US/10/252,256  
PRIOR FILING DATE: 2002-09-23  
PRIOR APPLICATION NUMBER: 60/165,334  
PRIOR FILING DATE: 1999-11-12  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 3  
LENGTH: 21  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: trfLunaasin  
US-10-252-256-3

Query Match  
Best Local Similarity 1.8%; Score 8; DB 9; Length 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 436 GDDDDDD 443  
Db 12 GDDDDDD 19

RESULT 5  
US-09-864-761-36493  
Sequence 36493, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Neomica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 36493  
LENGTH: 22  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC011235.2  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.97  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1  
US-09-864-761-36493

Query Match  
Best Local Similarity 1.8%; Score 8; DB 10; Length 22;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDD 444  
Db 7 DDDDDDD 14

RESULT 6  
US-10-252-256-5  
Sequence 5, Application US/10252256  
Publication No. US2003002765A1  
GENERAL INFORMATION:  
APPLICANT: GALVEZ, ALFREDO F.

```

1  TITLE OF INVENTION: THERAPEUTIC PEPTIDES HAVING A MOTIF THAT BINDS
2  TITLE OF INVENTION: SPECIFICALLY TO NON-ACETYLATED H3 AND H4 HISTONES FOR
3  TITLE OF INVENTION: CANCER THERAPY
4  FILE REFERENCE: 3729 . 02
5  CURRENT APPLICATION NUMBER: US/10/252,256
6  CURRENT FILING DATE: 2002-09-23
7  PRIOR APPLICATION NUMBER: 60/165,334
8  PRIOR FILING DATE: 1999-11-12
9  NUMBER OF SEQ ID NOS: 15
10 SOFTWARE: PatentIn Ver. 2.1
11 SEQ ID NO 5
12 LENGTH: 25
13 TYPE: PRT
14 ORGANISM: Glycine max
15 FEATURE:
16 OTHER INFORMATION: NLS-ctrlnasin
17 US-10-252-256-5

```

Query Match	1.8%	Score 8;	DB 9;	Length 25;
Best Similarity	100.0%	Pred. No. 2.9;		
Matches	8;	Conservative	0;	Indels 0; Gaps 0;

QY	436	GDDDDDD	443
Db	16	GDDDDDD	23

RESULT 7  
US-09-864-761-35294

GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David R.

; APPLICANT: Chen, wensheng  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 ; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23

PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: IIS 60/207.456

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR APPLICATION NUMBER: 05 24205:0  
PRIOR FILING DATE: 2000-10-04

PRIOR FILING DATE: 2000-09-2/  
PRIOR APPLICATION NUMBER: PCT/US01/0066

PRIOR APPLICATION NUMBER: PCT/US01/0066  
PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/006

PRIOR APPLICATION NUMBER: PCT/US01/006  
PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

;  
PRIOR FILING DATE: 2001-01-30  
;  
PRIOR APPLICATION NUMBER: PCT/US01/006  
;

PRIOR APPLICATION NUMBER: PCI/US01/006  
PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,688

```

1 Prior Application Number: US 09/608,408
2
3 Prior Filing Date: 2000-06-30
4
5 Prior Application Number: US 09/774,203
6
7 Prior Filing Date: 2001-01-29
8
9 Number of Seq. ID Nos: 49117
10
11 Software: Rnmax Sequence Listing Engine vers. 1.1.1
12
13 Seq ID No 35294

```

ORGANISM: Homo sapiens

OTHER INFORMATION: MAP TO AC011233.1  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3

```

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.0
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.93

```

OTHER INFORMATION:	EXPRESSED IN FETAL LIVER, SIGNAL = 2
OTHER INFORMATION:	EXPRESSED IN ADULT LIVER, SIGNAL = 1.1

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4  
EXPRESSED IN BONE MARROW, SIGNAL = 1.6

08-05-2014 02:24  
1 0%, score 8: DR 10: Length 31

Best Local Similarity	100.00%	100.00%	100.00%
Matches	8	Conservative	0
Mismatches	0	Indels	0
Gaps	0		

437 DDDDDDDG 44  
| | | | |  
QY

RESULT 8  
US-09-864-761-37113  
: Sequence 37113. Application US/09864761

Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; ADDITIONAL: Penn Sharron G

APPLICANT: Hanzel, David H

TITLE OF INVENTION:	HUMAN GENOMED-DERIVED GENES
TITLE OF INVENTION:	GENE EXPRESSION ANALYSIS BY MICROARRAY

;; CURRENT APPLICATION NUMBER: 05/09/864, 18  
;; CURRENT FILING DATE: 2001-05-23

PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03

PRIOR FILING DATE: 2000-10-04  
PRIORITY DOCUMENT NUMBER: IIS 60/236.359

; PRIOR FILING DATE: 2000 02 20  
 ; PRIOR APPLICATION NUMBER: PCT/US01/006666

-; PRIOR APPLICATION NUMBER: FC1/0301/0000  
 ; PRIOR FILING DATE: 2001-01-30

;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR APPLICATION NUMBER: PCT/US01/00662  
FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 37113  
LENGTH: 42  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AJ229043.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 9.4  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 8.3  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.5  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.5  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 8.8  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 9.1  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 8.3  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 8.9  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.9  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.9  
OTHER INFORMATION: EST\_HUMAN HIT: AMB63068.1, EVALUATE 3.00e-07  
US-09-864-761-37113

Query Match  
Best Local Similarity 1.8%; Score 8; DB 10; Length 42;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
DB 14 DDDDDDDG 21

RESULT 9  
US-10-252-256-1  
Sequence 1, Application US/10252256  
Publicatidn No. US20030027765A1  
GENERAL INFORMATION:  
APPLICANT: GALVEZ, ALFREDO F.  
TITLE OF INVENTION: THERAPEUTIC PEPTIDES HAVING A MOTIF THAT BINDS SPECIFICALLY TO NON-ACETYLATED H3 AND H4 HISTONES FOR CANCER THERAPY  
FILE REFERENCE: 3729.02  
CURRENT APPLICATION NUMBER: US/10/252,256  
CURRENT FILING DATE: 2002-09-23  
PRIOR APPLICATION NUMBER: 60/165,334  
PRIOR FILING DATE: 1999-11-12  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 43  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Lunasin  
US-10-252-256-1

Query Match  
Best Local Similarity 1.8%; Score 8; DB 9; Length 43;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDD 443  
DB 34 GDDDDDDD 41

RESULT 10  
US-09-836-392-34  
Sequence 34, Application US/09836392  
Patent No.: US20020173458A1  
GENERAL INFORMATION:  
APPLICANT: Ruben et al.  
TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptides, Antibodies  
FILE REFERENCE: PTO2001  
CURRENT APPLICATION NUMBER: US/09/836,392  
CURRENT FILING DATE: 2001-04-18  
PRIOR APPLICATION NUMBER: PCT/US00/28066  
PRIOR FILING DATE: 2000-10-11  
PRIOR APPLICATION NUMBER: 60/159,542  
PRIOR FILING DATE: 1999-10-15  
PRIOR APPLICATION NUMBER: 60/165,914  
PRIOR FILING DATE: 1999-11-17  
PRIOR APPLICATION NUMBER: 60/189,027  
PRIOR FILING DATE: 2000-03-14  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 34  
LENGTH: 49  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-836-392-34

Query Match  
Best Local Similarity 1.8%; Score 8; DB 9; Length 49;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153  
DB 30 VLKICDFG 37

RESULT 11  
US-09-809-545A-6  
Sequence 6, Application US/09809545A  
Patent No. US20020110804A1  
GENERAL INFORMATION:  
APPLICANT: Stanton, Lawrence W.  
APPLICANT: White, R. Tyler  
TITLE OF INVENTION: SECRETED FACTORS  
FILE REFERENCE: SCIOS.017A  
CURRENT APPLICATION NUMBER: US/09/809,545A  
CURRENT FILING DATE: 2001-03-14  
NUMBER OF SEQ ID NOS: 84  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 236  
TYPE: PRT  
ORGANISM: Rattus norvegicus  
US-09-809-545A-6

Query Match  
Best Local Similarity 1.8%; Score 8; DB 10; Length 236;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDDD 443  
DB 41 GDDDDDDD 48

RESULT 12  
US-09-801-368-188  
Sequence 188, Application US/09801368  
Patent No. US20020128250A1  
GENERAL INFORMATION:  
APPLICANT: Busby, Robert  
APPLICANT: Call, Brian  
APPLICANT: Hecht, Peter  
APPLICANT: Holtzman, Doug

APPLICANT: Madden, Kevin  
APPLICANT: Maxon, Mary  
APPLICANT: Milne, Todd  
APPLICANT: No. US20020128250A1man, Thea  
APPLICANT: Royer, John  
APPLICANT: Salama, Sofie  
APPLICANT: Sherman, Amir  
APPLICANT: Silva, Jeff  
APPLICANT: Summers, Eric  
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi  
FILE REFERENCE: 109272.147  
CURRENT APPLICATION NUMBER: US/09/801,368  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 09/487,558  
PRIOR FILING DATE: 2000-01-19  
PRIOR APPLICATION NUMBER: US 60/160,587  
NUMBER OF SEQ ID NOS: 440  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 188  
LENGTH: 286  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-801-368-188

Query Match 1.8%; Score 8; DB 10; Length 286;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
DB 108 GDDDDDD 115

RESULT 13  
US-09-862-027-18  
Sequence 18, Application US/09862027  
Patent No. US20020142428A1  
GENERAL INFORMATION:  
APPLICANT: Hodge, Martin R.  
TITLE OF INVENTION: No. US20020142428A1 Kinases and Uses Thereof  
FILE REFERENCE: 35800/234862  
CURRENT APPLICATION NUMBER: US/09/862,027  
CURRENT FILING DATE: 2001-05-21  
PRIOR APPLICATION NUMBER: US 09/345,473  
PRIOR FILING DATE: 1999-06-30  
NUMBER OF SEQ ID NOS: 82  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 18  
LENGTH: 328  
TYPE: PRT  
ORGANISM: C. elegans  
US-09-862-027-18

Query Match 1.8%; Score 8; DB 10; Length 328;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 EKEAEIIS 58  
DB 90 EKEAEIIS 97

RESULT 14  
US-09-291-417-5  
Sequence 5, Application US/09291417A  
Publication No. US20030050230A1  
GENERAL INFORMATION:  
APPLICANT: FLOWMAN, GREGORY  
APPLICANT: MARTINEZ, RICARDO  
APPLICANT: WHITE, DAVID  
TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES  
FILE REFERENCE: 240/300

CURRENT APPLICATION NUMBER: US/09/291,417A  
CURRENT FILING DATE: 1999-04-13  
EARLIER APPLICATION NUMBER: US 60/081,784  
EARLIER FILING DATE: 1998-04-14  
NUMBER OF SEQ ID NOS: 147  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 5  
LENGTH: 416  
TYPE: PRT  
ORGANISM: Mammalian (Human) STUK2  
US-09-291-417-5

Query Match 1.8%; Score 8; DB 9; Length 416;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIO 179  
DB 185 WMAPEVIO 192

RESULT 15  
US-09-862-027-14  
Sequence 14, Application US/09862027  
Patent No. US20020142428A1  
GENERAL INFORMATION:  
APPLICANT: Hodge, Martin R.  
TITLE OF INVENTION: No. US20020142428A1 Kinases and Uses Thereof  
FILE REFERENCE: 35800/234862  
CURRENT APPLICATION NUMBER: US/09/862,027  
CURRENT FILING DATE: 2001-05-21  
PRIOR APPLICATION NUMBER: US 09/345,473  
PRIOR FILING DATE: 1999-06-30  
NUMBER OF SEQ ID NOS: 82  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14  
LENGTH: 416  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-862-027-14

Query Match 1.8%; Score 8; DB 10; Length 416;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIO 179  
DB 185 WMAPEVIO 192

RESULT 16  
US-09-810-808-8  
Sequence 8, Application US/09810808  
Patent No. US20020042114A1  
GENERAL INFORMATION:  
APPLICANT: Au-Young, Janice  
Guegler, Karl U.  
Hawkins, Phillip R.  
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/810,808  
FILING DATE: 15-Mar-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/541,228  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0118 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 487 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: Genbank  
CLONE: 1117791  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-09-810-808-8

Query Match 1.8%; Score 8; DB 10; Length 487;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179  
DB 190 WMAPEVIQ 197

RESULT 17  
US-10-158-895-4  
Sequence 4, Application US/10158895  
Patent No. US20020155624A1  
GENERAL INFORMATION:  
APPLICANT: ONO, KOICHIRO  
APPLICANT: OHTOMO, TOSHIHIRO  
APPLICANT: TSUCHIYA, MASAYUKI  
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
FILE REFERENCE: 053466/0278  
CURRENT APPLICATION NUMBER: US/10/158,895  
CURRENT FILING DATE: 2002-06-03  
PRIOR APPLICATION NUMBER: US/09/529,279  
PRIOR FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: PCT/JP98/04796  
PRIOR FILING DATE: 1998-10-22  
PRIOR APPLICATION NUMBER: JP 9/290188  
PRIOR FILING DATE: 1997-10-22  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 579  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-158-895-4

Query Match 1.8%; Score 8; DB 9; Length 579;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153  
DB 170 VLKICDFG 177

RESULT 18  
US-10-158-895-15  
Sequence 15, Application US/10158895  
Patent No. US20020155624A1

GENERAL INFORMATION:  
APPLICANT: ONO, KOICHIRO  
APPLICANT: OHTOMO, TOSHIHIRO  
APPLICANT: TSUCHIYA, MASAYUKI  
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
FILE REFERENCE: 053466/0278  
CURRENT APPLICATION NUMBER: US/10/158,895  
CURRENT FILING DATE: 2002-06-03  
PRIOR APPLICATION NUMBER: US/09/529,279  
PRIOR FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: PCT/JP98/04796  
PRIOR FILING DATE: 1998-10-22  
PRIOR APPLICATION NUMBER: JP 9/290188  
PRIOR FILING DATE: 1997-10-22  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 15  
LENGTH: 590  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-158-895-15

Query Match 1.8%; Score 8; DB 9; Length 590;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153  
DB 170 VLKICDFG 177

RESULT 19  
US-10-143-133-2  
Sequence 2, Application US/10143133  
Publication No. US20020197658A1  
GENERAL INFORMATION:  
APPLICANT: Yoganathan, Thillainathan  
APPLICANT: Delaney, Allen  
TITLE OF INVENTION: Cancer Associated Protein Kinase and Its Use  
FILE REFERENCE: KINE-023  
CURRENT APPLICATION NUMBER: US/10/143,133  
CURRENT FILING DATE: 2002-05-09  
PRIOR APPLICATION NUMBER: 60/290,555  
PRIOR FILING DATE: 2001-05-10  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 847  
TYPE: PRT  
ORGANISM: Homo sapien  
US-10-143-133-2

Query Match 1.8%; Score 8; DB 9; Length 847;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VIHRLDKS 136  
DB 237 VIHRLDKS 244

RESULT 20  
US-09-882-291-22  
Sequence 22, Application US/09882291  
Publication No. US20030040472A1  
GENERAL INFORMATION:  
APPLICANT: Zealand Pharmaceuticals A/S  
TITLE OF INVENTION: No. US20030040472A1 Peptide Conjugates  
FILE REFERENCE: 007-2001  
CURRENT APPLICATION NUMBER: US/09/882,291  
CURRENT FILING DATE: 2001-06-15  
NUMBER OF SEQ ID NOS: 77  
SOFTWARE: PatentIn version 3.1

SEQ ID NO 22  
LENGTH: 8  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide sequence  
US-09-882-291-22

Query Match 1.5%; Score 7; DB 9; Length 8;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

OY 437 DDDDDD 443  
DB 2 DDDDDD 8

## RESULT 21

US-09-882-291-29  
Sequence 29, Application US/09882291  
Publication No. US20030040472A1

## GENERAL INFORMATION:

APPLICANT: Zealand Pharmaceuticals A/S  
TITLE OF INVENTION: No. US20030040472A1 Peptide Conjugates  
FILE REFERENCE: 007-2001  
CURRENT APPLICATION NUMBER: US/09/882,291  
CURRENT FILING DATE: 2001-06-15  
NUMBER OF SEQ ID NOS: 77  
SOFTWARE: Patentin version 3.1  
SEQ ID NO 29  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide sequence  
US-09-882-291-29

Query Match 1.5%; Score 7; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDD 443  
DB 3 DDDDDD 9

## RESULT 22

US-09-876-904A-173  
Sequence 173, Application US/09876904A  
Publication No. US20030072794A1

## GENERAL INFORMATION:

APPLICANT: BOULIVAS, TENI  
TITLE OF INVENTION: ENCAPSULATION OF PLASMID DNA (LIGASES TM) AND THERAPEUTIC  
TITLE OF INVENTION: AGENTS WITH NUCLEAR LOCALIZATION SIGNAL/FUSOGENIC PEPTIDE  
TITLE OF INVENTION: CONJUGATES INTO TARGETED LIPOSOME COMPLEXES  
FILE REFERENCE: TB-2002.00  
CURRENT APPLICATION NUMBER: US/09/876,904A  
CURRENT FILING DATE: 2001-06-08  
PRIOR APPLICATION NUMBER: US 60/210,925  
PRIOR FILING DATE: 2000-06-09  
NUMBER OF SEQ ID NOS: 629  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 173  
LENGTH: 12  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Karyophilic peptide  
US-09-876-904A-173

Query Match 1.5%; Score 7; DB 9; Length 12;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 129 VTHRLK 135  
DB 5 VTHRLK 11

## RESULT 23

US-09-911-569-80

Sequence 80, Application US/09911569  
Publication No. US20030069173A1

## GENERAL INFORMATION:

APPLICANT: HAMLEY-NELSON, PAMELA  
LAN, JIANQING  
SHIH, POJEN  
JESSE, JOEL A.  
SCHIFFERLI, KEVIN P.  
GEBREYEHU, GULIAT  
TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS  
NUMBER OF SEQUENCES: 120  
CORRESPONDENCE ADDRESS:  
ADDRESSER: GREENLEE, WINNER & SULLIVAN  
STREET: 5370 MANHATTAN CIRCLE, SUITE 201  
CITY: BOULDER  
STATE: CO  
COUNTRY: US  
ZIP: 80303

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/911,569  
FILING DATE: 23-Jul-2001  
CLASSIFICATION: <Unknown>

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/039,780  
FILING DATE: 16-MAR-1998

## ATTORNEY/AGENT INFORMATION:

NAME: SULLIVAN, SALLY A.  
REGISTRATION NUMBER: 32,064  
REFERENCE/DOCKET NUMBER: 32-95D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303)499-8080  
TELEFAX: (303)499-8089

## INFORMATION FOR SEQ ID NO: 80:

SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: Modified-site

LOCATION: 14  
OTHER INFORMATION: /product= "OTHER"  
/note= "G AT POSITION 14 CAN BE ABSENT"

## SEQUENCE DESCRIPTION: SEQ ID NO: 80:

US-09-911-569-80  
Query Match 1.5%; Score 7; DB 9; Length 14;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 438 DDDDDG 444  
DB 8 DDDDDG 14

## RESULT 24

US-09-864-761-38908





Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443  
Db 3 DDDDDD 9

```
RESULT 26
US-09-864-761-34620
; Sequence 34620, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34620
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009491.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN PLACENTAR, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
```

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3  
US-09-864-761-34620

Query Match 1.5%; Score 7; DB 10; Length 24;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443  
Db 5 DDDDDD 11

```
RESULT 27
US-09-864-761-36142
; Sequence 36142, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36142
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007159.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.7
```

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.2  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.1  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 6.2  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2  
US-09-864-761-36142

Query Match 1.5%; Score 7; DB 10; Length 24;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDD 443  
Db 1 DDDDDDD 7

## RESULT 28

Sequence 35447, Application US/09864761  
Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: Aecm1ca-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

PRIOR FILING DATE: 2001-05-23

PRIOR FILING DATE: 2000-02-04

PRIOR FILING DATE: 2000-02-04

PRIOR FILING DATE: 2000-05-26

PRIOR FILING DATE: 2000-05-26

PRIOR FILING DATE: 2000-08-03

PRIOR FILING DATE: 2000-10-04

PRIOR FILING DATE: 2000-09-27

PRIOR FILING DATE: 2000-09-27

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AL031661.16

OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 2.5

OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.1

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2

OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3

OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7

OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.3

US-09-864-761-35447

Query Match 1.5%; Score 7; DB 10; Length 26;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDD 443  
Db 3 DDDDDDD 9

## RESULT 29

Sequence 39392

Sequence 39392, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: Aecm1ca-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

PRIOR FILING DATE: 2001-05-23

PRIOR FILING DATE: 2000-02-04

PRIOR FILING DATE: 2000-02-04

PRIOR FILING DATE: 2000-05-26

PRIOR FILING DATE: 2000-05-26

PRIOR FILING DATE: 2000-08-03

PRIOR FILING DATE: 2000-10-04

PRIOR FILING DATE: 2000-09-27

PRIOR FILING DATE: 2000-09-27

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

PRIOR FILING DATE: 2001-01-30

NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 39392  
LENGTH: 27  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC005959.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN RETAL LIVER, SIGNAL = 1.7  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7  
OTHER INFORMATION: EST\_HUMAN HIT: T95763.1, EVALUE 2.10e+00  
US-09-864-761-39392

Query Match 1.5%; Score 7; DB 10; Length 27;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
DB 3 DDDDDDD 9

RESULT 30  
US-09-911-569-87  
Sequence 87, Application US/09911569  
Publication No. US20030069173A1  
GENERAL INFORMATION:  
APPLICANT: HAMLEY-NELSON, PAMELA  
LAN, JIANQING  
SHIH, POJEN  
JESSE, JOEL A.  
SCHIFFERLI, KEVIN P.  
GEBREYEHU, GULILAT  
TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSECTIONS  
NUMBER OF SEQUENCES: 120  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GREENLEE, WINNER & SULLIVAN  
STREET: 5370 MANHATTAN CIRCLE, SUITE 201  
CITY: BOULDER  
STATE: CO  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/911,569  
FILING DATE: 23-Jul-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/039,780  
FILING DATE: 16-MAR-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: SULLIVAN, SALLY A.  
REGISTRATION NUMBER: 32,064  
REFERENCE/DOCKET NUMBER: 32-95D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303)499-8080  
TELEFAX: (303)499-8089  
INFORMATION FOR SEQ ID NO: 87:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 30  
OTHER INFORMATION: /product= "OTHER"  
/note= "G AT POSITION 30 CAN BE ABSENT"  
SEQUENCE DESCRIPTION: SEQ ID NO: 87:  
US-09-911-569-87

Query Match 1.5%; Score 7; DB 9; Length 30;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 438 DDDDDDG 444  
DB 24 DDDDDDG 30

RESULT 31  
US-09-864-761-33555  
Sequence 33555, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecm1ca-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 33555

```
LENGTH: 31
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AP000032.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.2
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4
US-09-864-761-33555

Query Match
Best Local Similarity 1.5% Score 7; DB 10; Length 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 437 DDDDDDD 443
Db 5 DDDDDDD 11

RESULT 32
US-09-864-761-44046
Sequence 44046, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR APPLICATION NUMBER: US 09/608,408
```

```
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 44046
LENGTH: 34
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC002478.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.66
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.74
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.73
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.88
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.65
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.78
US-09-864-761-44046

Query Match
Best Local Similarity 1.5% Score 7; DB 10; Length 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 437 DDDDDDD 443
Db 4 DDDDDDD 10

RESULT 33
US-09-864-761-35933
Sequence 35933, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
```

```
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 35933
;; LENGTH: 36
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC010792.1
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.3
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.6
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.4
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 2.3
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
US-09-864-761-35933

Query Match          1.5%; Score 7; DB 10; Length 36;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
DB 9 DDDDDDD 15

RESULT 34
US-09-864-761-43116
;; Sequence 43116, Application US/09864761
;; Patent No. US2002048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;; FILE REFERENCE: Aecm1ca-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
```

```
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 43116
;; LENGTH: 37
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL008631.1
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
;; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.4
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.95
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.6
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
US-09-864-761-43116

Query Match          1.5%; Score 7; DB 10; Length 37;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
DB 3 DDDDDDD 9

RESULT 35
US-10-252-256-2
;; Sequence 2, Application US/10252256
;; Publication No. US20030027765A1
;; GENERAL INFORMATION:
;; APPLICANT: GALVEZ, ALFREDO F.
;; TITLE OF INVENTION: THERAPEUTIC PEPTIDES HAVING A MOTIF THAT BINDS
;; TITLE OF INVENTION: SPECIFICALLY TO NON-ACTYLATED H3 AND H4 HISTONES FOR
;; TITLE OF INVENTION: CANCER THERAPY
;; FILE REFERENCE: 3729.02
;; CURRENT APPLICATION NUMBER: US/10/252,256
;; CURRENT FILING DATE: 2002-09-23
;; PRIOR APPLICATION NUMBER: 60/165,334
;; PRIOR FILING DATE: 1999-11-12
;; NUMBER OF SEQ ID NOS: 15
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 2
;; LENGTH: 40
;; TYPE: PRT
;; ORGANISM: Glycine max
;; FEATURE:
;; OTHER INFORMATION: Lunasin-GRG
US-10-252-256-2

Query Match          1.5%; Score 7; DB 9; Length 40;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
DB 32 DDDDDDD 38

RESULT 36
US-09-864-761-39102
```

```

: Sequence 39102, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharron G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Acomica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 39102
: LENGTH: 40
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO ACO05873.2
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.3
: OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 2.1
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
: US-09-864-761-39102
:
: Query Match 1.5%; Score 7; DB 10; Length 40;
: Best Local Similarity 100.0%; Pred. No. 36;
: Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: 437 DDDDDD 443
: 6 DDDDDD 12

```

RESULT 37.  
 US-09-864-761-35018  
 ; Sequence 35018, Application US/09864761  
 ; Patent No. US20020048763A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharron G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanzel, David K.  
 ; APPLICANT: Chen, Wensheng  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 ; FILE REFERENCE: Aecomica-X-1  
 ; CURRENT APPLICATION NUMBER: US/09/864,761  
 ; CURRENT FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: US 60/180,312  
 ; PRIOR FILING DATE: 2000-02-04  
 ; PRIOR APPLICATION NUMBER: US 60/207,456  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/632,366  
 ; PRIOR FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: GB 24263.6  
 ; PRIOR FILING DATE: 2000-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/236,359  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: US 60/234,687  
 ; PRIOR FILING DATE: 2000-09-21  
 ; PRIOR APPLICATION NUMBER: US 09/608,408  
 ; PRIOR FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: US 09/774,203  
 ; PRIOR FILING DATE: 2001-01-29  
 ; NUMBER OF SEQ ID NOS: 49117  
 ; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
 ; SEQ ID NO 35018  
 ; LENGTH: 42  
 ; TYPE: prt  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: MAP TO AL049553.14  
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.3  
 ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4  
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6  
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8  
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2  
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9  
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4  
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.5  
 ; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.4  
 ; US-09-864-761-35018

Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDD 443  
|||||  
Db 5 DDDDDDD 11

```

RESULT 38
US-09-864-761-35629
; Sequence 35629, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmca-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35629
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009492.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN PETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6

```

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.3  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.98  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2  
US-09-864-761-35629

Query Match 1.5%; Score 7; DB 10; Length 42;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 436 GDDDDDD 442  
|||||  
Db 21 GDDDDDD 27

```

RESULT 39
US-09-864-761-39869
; Sequence 39869, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmca-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 39869
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004253.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2

```



OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.3  
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 2.2  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1  
US-09-864-761-39869

Query Match 1.5%; Score 7; DB 10; Length 51;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443  
Db 1 DDDDDD 7

## RESULT 40

US-09-864-761-39079  
Sequence 39079, Application US/09864761  
Patent No. US20020048763A1

## GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aegmics-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 39079  
LENGTH: 61  
TYPE: PRT

ORGANISM: Homo sapiens  
FEATURE:

OTHER INFORMATION: MAP TO AC004912.1  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2  
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7  
OTHER INFORMATION: SWISSPROT HIT: Q09322, EVALU 3.50e+00  
US-09-864-761-39079

Query Match 1.5%; Score 7; DB 10; Length 61;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443  
Db 6 DDDDDD 12

## RESULT 41

US-10-028-247-2  
Sequence 2, Application US/10028247  
Patent No. US20020150594A1

## GENERAL INFORMATION:

APPLICANT: Goldman, Stanley  
APPLICANT: Lathrop, Stephanie J.  
APPLICANT: Longchamp, Pascal F.  
APPLICANT: Whalen, Robert G.  
APPLICANT: Maxygen, Inc.  
TITLE OF INVENTION: Methods and Compositions for Developing Spore Display  
FILE REFERENCE: 18097A-033520US  
CURRENT APPLICATION NUMBER: US/10/028,247  
CURRENT FILING DATE: 2001-12-19  
PRIOR APPLICATION NUMBER: US 60/214,161  
PRIOR FILING DATE: 2000-06-26  
PRIOR APPLICATION NUMBER: US 09/892,208  
PRIOR FILING DATE: 2001-06-26  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 86  
TYPE: PRT  
ORGANISM: Bacillus subtilis  
FEATURE:  
OTHER INFORMATION: ColC27 including Hail epitope region  
US-10-028-247-2

Query Match 1.5%; Score 7; DB 12; Length 86;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 YASLGS 90  
Db 40 YASLGS 46

## RESULT 42

US-09-873-880-4  
Sequence 4, Application US/09873880  
Patent No. US20020123118A1

## GENERAL INFORMATION:

APPLICANT: Sewalt, Vincent  
APPLICANT: Falco, S. Carl  
APPLICANT: Allen, Stephen M.  
TITLE OF INVENTION: GLYCINE METABOLISM ENZYMES  
FILE REFERENCE: BB1192 US CIP  
CURRENT APPLICATION NUMBER: US/09/873,880  
CURRENT FILING DATE: 2001-06-04

```
;; PRIOR APPLICATION NUMBER: 09/363,321
;; PRIOR FILING DATE: July 28, 1999
;; PRIOR APPLICATION NUMBER: 60/094,839
;; PRIOR FILING DATE: July 31, 1998
;; NUMBER OF SEQ ID NOS: 42
;; SOFTWARE: Microsoft Office 97
;; SEQ ID NO 4
;; LENGTH: 91
;; TYPE: PRT
;; ORGANISM: Zea mays
US-09-873-880-4

Query Match      1.5%; Score 7; DB 10; Length 91;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 364 SNGEGHG 370
Db 4 SNGEGHG 10

RESULT 43
US-09-864-761-44209
;; Sequence 44209, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: Aecmca-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263,6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
```

```
;; SEQ ID NO 44209
;; LENGTH: 102
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC005668.1
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.55
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
US-09-864-761-44209

Query Match      1.5%; Score 7; DB 10; Length 102;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
Db 35 DDDDDDD 41

RESULT 44
US-09-864-761-42215
;; Sequence 42215, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: Aecmca-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263,6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
```

NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1  
SEQ ID NO 42215  
LENGTH: 110  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC018801.2.  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.2  
OTHER INFORMATION: EXPRESSED IN HBLA, SIGNAL = 4.6  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.8  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 7.3  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.5  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.4  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.2  
OTHER INFORMATION: EST HUMAN HIT: BF507876.1, EVALUATE 7.00e-55  
OTHER INFORMATION: SWISSPROT HIT: P34657, EVALUATE 3.00e-21  
US-09-864-761-42215

Query Match 1.5%; Score 7; DB 10; Length 110;  
Best Local Similarity 100.0%; Pred. No. 86;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 437 DDDDDDD 443  
Db 103 DDDDDDD 109

RESULT 45  
US-10-101-464A-662  
Sequence 662, Application US/10101464A  
Publication No. US20030046728A1  
GENERAL INFORMATION:  
APPLICANT: Strabala, Timothy  
APPLICANT: Nieuwenhuizen, Nicolaas  
TITLE OF INVENTION: Compositions isolated from plant cells  
TITLE OF INVENTION: and their use in the modification of plant cell signaling  
FILE REFERENCE: 11000.1020c2  
CURRENT APPLICATION NUMBER: US/101/101,464A  
PRIOR FILING DATE: 2002-03-18  
PRIOR APPLICATION NUMBER: 09/704,302  
PRIOR FILING DATE: 2000-11-01  
PRIOR APPLICATION NUMBER: 09/228,986  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/162,866  
PRIOR FILING DATE: 1999-11-01  
PRIOR APPLICATION NUMBER: PCT/US00/00724  
PRIOR FILING DATE: 2000-01-11  
NUMBER OF SEQ ID NOS: 989  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 662  
LENGTH: 154  
TYPE: PRT  
ORGANISM: Pinus radiata  
US-10-101-464A-662

Query Match 1.5%; Score 7; DB 9; Length 154;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 25 GSGGSVY 31  
Db 27 GSGGSVY 33

RESULT 46  
US-09-864-761-33375  
Sequence 33375, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecmca-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
PRIOR FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1  
SEQ ID NO 33375  
LENGTH: 154  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC005921.3  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.9  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN HEILA, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6.9  
OTHER INFORMATION: SWISSPROT HIT: Q02926, EVALUATE 6.30e-01  
OTHER INFORMATION: EST HUMAN HIT: BE741337.1, EVALUATE 2.00e-54  
OTHER INFORMATION: EST\_HUMAN HIT: D78804.1, EVALUATE 6.00e-35  
US-09-864-761-33375

Query Match 1.5%; Score 7; DB 10; Length 154;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 437 DDDDDDD 443  
Db 103 DDDDDDD 109

Db 4 DDDDDDD 10

## RESULT 47

US-09-864-761-43542  
; Sequence 43542, Application US/09864761  
; Patent No. US20020048763A1

## GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aomic-X-1  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

CURRENT FILING DATE: 2001-05-23, 761

PRIOR FILING DATE: 2000-02-04, 312

PRIOR FILING DATE: 2000-02-04, 456

PRIOR FILING DATE: 2000-05-26, 366

PRIOR FILING DATE: 2000-08-03, 6

PRIOR FILING DATE: 2000-10-04, 359

PRIOR FILING DATE: 2000-09-27, 666

PRIOR FILING DATE: 2001-01-30, 667

PRIOR FILING DATE: 2001-01-30, 664

PRIOR FILING DATE: 2001-01-30, 669

PRIOR FILING DATE: 2001-01-30, 665

PRIOR FILING DATE: 2001-01-30, 668

PRIOR FILING DATE: 2001-01-30, 663

PRIOR FILING DATE: 2001-01-30, 662

PRIOR FILING DATE: 2001-01-30, 661

PRIOR FILING DATE: 2001-01-30, 670

PRIOR FILING DATE: 2001-01-30, 687

PRIOR FILING DATE: 2000-09-21, 408

PRIOR FILING DATE: 2000-06-30, 203

PRIOR FILING DATE: 2001-01-29, 117

NUMBER OF SEQ ID NOS: 49117

SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

SEQ ID NO 43542

LENGTH: 161

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE: MAP TO AC004816.1

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.92

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1

OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.88

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1

OTHER INFORMATION: EST\_HUMAN HIT: AM452931.1, EVALU 4.00e-53

US-09-864-761-43542

Query Match 1.5%; Score 7; DB 10; Length 161;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 130 IHRDLKS 136  
Db 129 IHRDLKS 135

## RESULT 48

US-10-101-464A-681  
; Sequence 681, Application US/10101464A  
; Publication No. US20030046728A1

## GENERAL INFORMATION:

APPLICANT: Strabala, Timothy  
APPLICANT: Nieuwenhuizen, NicolaasAPPLICANT: Higgins, Colleen M.  
TITLE OF INVENTION: Compositions Isolated from Plant CellsFILE REFERENCE: 11000.1020C2  
TITLE OF INVENTION: and their use in the Modification of Plant Cell Signaling

CURRENT FILING DATE: 2002-03-18, 464A

PRIOR FILING DATE: 2000-11-01, 302

PRIOR FILING DATE: 2000-11-01, 986

PRIOR FILING DATE: 1999-01-12, 866

PRIOR FILING DATE: 1999-11-01, 00724

PRIOR FILING DATE: 2000-01-11, 989

NUMBER OF SEQ ID NOS: 989

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 681

LENGTH: 163

TYPE: PRT

ORGANISM: Pinus radiata

US-10-101-464A-681

Query Match 1.5%; Score 7; DB 9; Length 163;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GSFGSVY 31

Db 99 GSFGSVY 105

RESULT 49

US-09-764-868-816

; Sequence 816, Application US/09764868

; Patent No. US20020168711A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PT232

CURRENT FILING DATE: 2001-01-17, 868

PRIOR FILING DATE: 2001-01-17, 1510

NUMBER OF SEQ ID NOS: 1510

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 816

LENGTH: 170

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE: NAME/KEY: SITE

LOCATION: (116)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-764-868-816

Query Match 1.5%; Score 7; DB 9; Length 170;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVI 178

|||||  
Db 94 WMAPEVI 100

## RESULT 50

US-09-764-868-845  
; Sequence 845, Application US/09764868  
; Patent No. US20020168711A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT232  
; CURRENT APPLICATION NUMBER: US/09/764,868  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1510  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 845  
; LENGTH: 170  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-868-845

Query Match 1.5%; Score 7; DB 9; Length 170;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 172 WMAPEVI 178  
|||||  
Db 96 WMAPEVI 102

## RESULT 51

US-09-764-891-4540  
; Sequence 4540, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4540  
; LENGTH: 170  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-891-4540

Query Match 1.5%; Score 7; DB 9; Length 170;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 172 WMAPEVI 178  
|||||  
Db 96 WMAPEVI 102

## RESULT 52

US-09-764-868-633  
; Sequence 633, Application US/09764868  
; Patent No. US20020168711A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT232  
; CURRENT APPLICATION NUMBER: US/09/764,868  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1510  
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 633  
; LENGTH: 174  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-868-633

Query Match 1.5%; Score 7; DB 9; Length 174;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 437 DDDDDDD 443  
|||||  
Db 165 DDDDDDD 171

## RESULT 53

US-09-764-868-833  
; Sequence 833, Application US/09764868  
; Patent No. US20020168711A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT232  
; CURRENT APPLICATION NUMBER: US/09/764,868  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1510  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 833  
; LENGTH: 177  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-868-833

Query Match 1.5%; Score 7; DB 9; Length 177;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 172 WMAPEVI 178  
|||||  
Db 29 WMAPEVI 35

## RESULT 54

US-09-764-891-5028  
; Sequence 5028, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5028  
; LENGTH: 177  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-891-5028

Query Match 1.5%; Score 7; DB 9; Length 177;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 172 WMAPEVI 178  
|||||  
Db 29 WMAPEVI 35

RESULT 55  
US-10-101-464A-540

```
; Sequence 540, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10101,464A
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 540
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-101-464A-540
```

```
Query Match 1.5%; Score 7; DB 9; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 129 VHRDLK 135
Db 140 VHRDLK 146
```

```
RESULT 56
US-09-738-626-4510
; Sequence 4510, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4510
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4510
```

```
Query Match 1.5%; Score 7; DB 9; Length 207;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 437 DDDDDD 443
Db 102 DDDDDD 108
```

```
RESULT 57
US-10-101-464A-702
; Sequence 702, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10101,464A
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 702
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Pinus radiata
US-10-101-464A-702
```

```
Query Match 1.5%; Score 7; DB 9; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 129 VHRDLK 135
Db 183 VHRDLK 189
```

```
RESULT 58
US-09-925-301-1424
; Sequence 1424, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1424
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (59)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (62)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (221)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
```

```
Query Match 1.5%; Score 7; DB 9; Length 238;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

US-09-925-301-1424

Query Match  
Best Local Similarity 100.0%; Score 7; DB 10; Length 244;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 TPLLLPL 334  
DB 15 TPLLLPL 21

RESULT 59

US-09-764-868-1063  
; Sequence 1063, Application US/09764868  
; Patent No. US2002016871A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT232  
; CURRENT APPLICATION NUMBER: US/09/764,868  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1510  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 1063  
; LENGTH: 250  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-764-868-1063

Query Match  
Best Local Similarity 100.0%; Score 7; DB 9; Length 250;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
DB 241 DDDDDDD 247

RESULT 60

US-09-955-999-77  
; Sequence 77, Application US/09955999  
; Publication No. US20030036505A1  
; GENERAL INFORMATION:  
; APPLICANT: Barash et al.  
; TITLE OF INVENTION: Signal Transduction Pathway Component Polynucleotides, Polypeptides  
; FILE REFERENCE: PT086P1  
; CURRENT APPLICATION NUMBER: US/09/955,999  
; CURRENT FILING DATE: 2001-09-20  
; PRIOR APPLICATION NUMBER: 60/234,997  
; PRIOR FILING DATE: 2000-09-25  
; NUMBER OF SEQ ID NOS: 139  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 77  
; LENGTH: 250  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-955-999-77

Query Match  
Best Local Similarity 100.0%; Score 7; DB 9; Length 250;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
DB 241 DDDDDDD 247

RESULT 61

US-09-898-837A-32  
; Sequence 32, Application US/09898837A  
; Publication No. US20030077697A1

GENERAL INFORMATION:

APPLICANT: Quinn, Kerry E.  
APPLICANT: Spylek, Kimberly A.  
APPLICANT: Majumder, Kuntud  
APPLICANT: Vernet, Corine  
APPLICANT: Herrmann, John L.  
APPLICANT: Burgess, Catherine  
APPLICANT: Fernandes, Elma  
APPLICANT: Taupier Jr., Raymond  
APPLICANT: Raestell, Luca  
APPLICANT: Curagen Corporation  
APPLICANT: Gerlach, Valerie L.  
APPLICANT: MacDougall, John R.  
TITLE OF INVENTION: NOVEL SERINE/THREONINE PROTEIN-KINASE LIKE PROTEINS AND  
TITLE OF INVENTION: NOVEL SERINE/THREONINE PROTEIN-KINASE LIKE PROTEINS AND  
FILE REFERENCE: 15966-598 CIP

CURRENT APPLICATION NUMBER: US/09/898,837A  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/165,986  
; PRIOR FILING DATE: 2001-07-03  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,839  
; PRIOR FILING DATE: 1999-11-17  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,839  
; PRIOR FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/195,637  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/197,080  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/232,677  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/181,347  
; PRIOR FILING DATE: 2000-02-09  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,195  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/215,906  
; PRIOR FILING DATE: 2000-07-03  
; PRIOR APPLICATION NUMBER: U.S.S.N. 09/715,427  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 32  
; LENGTH: 256  
; TYPE: PRT  
; ORGANISM: Saccharomyces cerevisiae  
US-09-898-837A-32

Query Match  
Best Local Similarity 100.0%; Score 7; DB 9; Length 256;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VHRDLK 135  
DB 119 VHRDLK 125

RESULT 62

US-09-949-842-14  
; Sequence 14, Application US/09949842  
; Patent No. US20020164692A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: PT047P1  
; FILE REFERENCE: Immune System-Related Polynucleotides, Polypeptides, and Antibodies  
; CURRENT APPLICATION NUMBER: US/09/949,842  
; CURRENT FILING DATE: 2001-09-02  
; PRIOR APPLICATION NUMBER: PCT/US01/07260  
; PRIOR FILING DATE: 2001-03-07  
; PRIOR APPLICATION NUMBER: 60/224,367  
; PRIOR FILING DATE: 2000-08-11  
; PRIOR APPLICATION NUMBER: 60/187,873  
; PRIOR FILING DATE: 2000-03-08  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 257

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-821
Query Match
Best Local Similarity 1.5%; Score 7; DB 9; Length 257;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443
Db 108 DDDDDD 114

RESULT 63
US-09-764-868-821
; Sequence 821, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 821
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (11)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (191)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (246)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (251)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (259)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (262)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (264)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-821

Query Match
Best Local Similarity 1.5%; Score 7; DB 9; Length 265;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVI 178
Db 98 WMAPEVI 104

RESULT 64
US-09-840-704-6
; Sequence 6, Application US/09840704
; Patent No. US20020122801A1
; GENERAL INFORMATION:
; APPLICANT: Dedhar, Shoukat
; APPLICANT: Hannigan, Greg
```

```
; TITLE OF INVENTION: Interglin-Linked Kinase and its Uses
; FILE REFERENCE: KIN-2CON
; CURRENT APPLICATION NUMBER: US/09/840,704
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 09/566,906
; PRIOR FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: US08/752,345
; PRIOR FILING DATE: 1996-11-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 271
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: Other
; LOCATION: (1)...(271)
US-09-840-704-6

Query Match
Best Local Similarity 1.5%; Score 7; DB 10; Length 271;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 IHRDLKS 136
Db 122 IHRDLKS 128

RESULT 65
US-10-274-409-5
; Sequence 5, Application US/10274409
; Publication No. US20030054529A1
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001161DIV
; CURRENT APPLICATION NUMBER: US/10/274,409
; CURRENT FILING DATE: 2002-10-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Dictyostelium discoideum
US-10-274-409-5

Query Match
Best Local Similarity 1.5%; Score 7; DB 9; Length 275;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVI 178
Db 184 WMAPEVI 190

RESULT 66
US-10-274-409-6
; Sequence 6, Application US/10274409
; Publication No. US20030054529A1
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001161DIV
; CURRENT APPLICATION NUMBER: US/10/274,409
; CURRENT FILING DATE: 2002-10-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 277
```



```

; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-274-409-6
Query Match
Best Local Similarity 1.5%; Score 7; DB 9; Length 277;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 172 WMAPEVI 178
Db 182 WMAPEVI 188

RESULT 67
US-10-274-409-7
; Sequence 7, Application US/10274409
; Publication No. US20030054529A1
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001161DIY
; CURRENT APPLICATION NUMBER: US/10/274,409
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FASTSEQ for windows Version 4.0
; SEQ ID NO: 7
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-274-409-7
Query Match
Best Local Similarity 1.5%; Score 7; DB 9; Length 277;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 172 WMAPEVI 178
Db 182 WMAPEVI 188

RESULT 68
US-08-910-386A-14
; Sequence 14, Application US/08910386A
; Patent No. US20020092041A1
; GENERAL INFORMATION:
; APPLICANT: Ronald, Pamela C.
; APPLICANT: Wang, Guo-Liang
; APPLICANT: Song, Wen-Yuang
; APPLICANT: Hulbert, Scot
; APPLICANT: Richter, Todd
; TITLE OF INVENTION: Procedures and Materials for Confering
; TITLE OF INVENTION: Disease Resistance in Plants
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,386A
; FILING DATE: 13-AUG-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Baastian, Kevin L.
```

```

; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-058950US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 280 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-910-386A-14
Query Match
Best Local Similarity 1.5%; Score 7; DB 8; Length 280;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 25 GSPGSHV 31
Db 143 GSPGSHV 149

RESULT 69
US-09-771-161A-106
; Sequence 106, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: Patentin version 3.0
; SEQ ID NO: 106
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-106
Query Match
Best Local Similarity 1.5%; Score 7; DB 10; Length 284;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 130 IHRDLKS 136
Db 276 IHRDLKS 282

RESULT 70
US-10-060-019-22
; Sequence 22, Application US/10060019
; Publication No. US20030003564A1
; GENERAL INFORMATION:
; APPLICANT: Tyers, Mike
; APPLICANT: Williams, Andrew
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING UBIQUITIN
; TITLE OF INVENTION: DEPENDENT PROTEOLYSIS
; FILE REFERENCE: 11757.10USU1
; CURRENT APPLICATION NUMBER: US/10/060,019
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: US/09/177,165
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/092,443
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: 60/063,254
; PRIOR FILING DATE: 1997-10-24
```

NUMBER OF SEQ ID NOS: 50  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 22  
LENGTH: 295  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-10-060-019-22

Query Match 1.5%; Score 7; DB 9; Length 295;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443  
DB 255 DDDDDD 261

## RESULT 71

US-09-991-211-5  
Sequence 5, Application US/09991211  
Patent No. US2002015685A1

GENERAL INFORMATION:  
APPLICANT: Pharmacia & Upjohn  
TITLE OF INVENTION: CRYSTALLIZATION AND STRUCTURE DETERMINATION OF  
TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS  
TITLE OF INVENTION: UDP-N-ACETYLENOLPYRUVYLGLUCOSAMINE REDUCTASE (S. aureus  
TITLE OF INVENTION: Murb)  
FILE REFERENCE: 268,6241 0101  
CURRENT APPLICATION NUMBER: US/09/991,211  
CURRENT FILING DATE: 2001-11-21  
PRIOR APPLICATION NUMBER: 09/632,947  
PRIOR FILING DATE: 2000-08-04  
PRIOR APPLICATION NUMBER: 60/147,164  
PRIOR FILING DATE: 1999-08-04  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 5  
LENGTH: 303  
TYPE: PRT  
ORGANISM: Bacillus subtilis  
US-09-991-211-5

Query Match 1.5%; Score 7; DB 9; Length 303;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 306 QELKERE 312  
DB 6 QELKERE 12

## RESULT 72

US-09-866-582-27  
Sequence 27, Application US/09666582  
Patent No. US20020127620A1  
GENERAL INFORMATION:  
APPLICANT: Witman, George B.  
APPLICANT: Pazour, Gregory J.  
APPLICANT: Rosenbaum, Joel L.  
APPLICANT: Cole, Douglas G.  
TITLE OF INVENTION: INTRAFAGELLAR TRANSPORT  
FILE REFERENCE: 07917-145001  
CURRENT APPLICATION NUMBER: US/09/866,582  
CURRENT FILING DATE: 2001-05-24  
PRIOR APPLICATION NUMBER: US 60/206,923  
PRIOR FILING DATE: 2000-05-24  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 27  
LENGTH: 304  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-866-582-27

Query Match 1.5%; Score 7; DB 10; Length 304;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443  
DB 34 DDDDDD 40

## RESULT 73

US-09-746-801A-61  
Sequence 61, Application US/09746801A  
Patent No. US2002008349A1  
GENERAL INFORMATION:  
APPLICANT: Wagner, et al.  
TITLE OF INVENTION: GENES REGULATING CIRCADIAN CLOCK FUNCTION AND PHOTOPERIODISM  
FILE REFERENCE: 1505-54357  
CURRENT APPLICATION NUMBER: US/09/746,801A  
CURRENT FILING DATE: 2000-12-20  
NUMBER OF SEQ ID NOS: 68  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 61  
LENGTH: 324  
TYPE: PRT  
ORGANISM: Brassica sp.

FEATURE: misc feature  
NAME/KEY: (188)..(188)  
OTHER INFORMATION: Xaa = uncertain amino acid residue  
NAME/KEY: misc feature  
LOCATION: (194)..(194)  
OTHER INFORMATION: Xaa = uncertain amino acid residue  
NAME/KEY: misc feature  
LOCATION: (196)..(196)  
OTHER INFORMATION: Xaa = uncertain amino acid residue  
NAME/KEY: misc feature  
LOCATION: (214)..(214)  
OTHER INFORMATION: Xaa = uncertain amino acid residue  
US-09-746-801A-61

Query Match 1.5%; Score 7; DB 10; Length 324;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443  
DB 162 DDDDDD 168

## RESULT 74

US-09-815-242-5783  
Sequence 5783, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari L.  
APPLICANT: Zykkind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727

Search completed: May 1, 2003, 21:01:50  
Job time : 54 secs

;; PRIOR FILING DATE: 2000-05-26  
;; PRIOR APPLICATION NUMBER: 60/242,578  
;; PRIOR FILING DATE: 2000-10-23  
;; PRIOR APPLICATION NUMBER: 60/253,625  
;; PRIOR FILING DATE: 2000-11-27  
;; PRIOR APPLICATION NUMBER: 60/257,931  
;; PRIOR FILING DATE: 2000-12-22  
;; PRIOR APPLICATION NUMBER: 60/269,308  
;; PRIOR FILING DATE: 2001-02-16  
;; NUMBER OF SEQ ID NOS: 14110  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO: 5783  
;; LENGTH: 325  
;; TYPE: PRT  
;; ORGANISM: Staphylococcus aureus  
US-09-815-242-5783

Query Match 1.5%; Score 7; DB 10; Length 325;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

57 LSVLSHR 63  
|||||  
Db 252 LSVLSHR 258

RESULT 75  
US-09-815-242-12693  
; Sequence 12693, Application US/09815242  
; Patent No. US2002061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zysek, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.01A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 12693  
; LENGTH: 325  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-09-815-242-12693

Query Match 1.5%; Score 7; DB 10; Length 325;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 LSVLSHR 63  
|||||  
Db 252 LSVLSHR 258



GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 20:52:23 ; Search time 19 Seconds  
(without alignments)  
704.602 Million cell updates/sec

Title: US-09-757-982-5

Perfect score: 455

Sequence: 1 MSSLGASFVQIKFDLQFFE.....GDDDDDDGEEEDNDMDNSE 455

Scoring table: OLIGO  
Gapop 60.0, Gapext 60.0

Searched: 262574 seqs, 29422922 residues

rd size: 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 100 summaries

Database: Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*

2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*

3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*

4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*

5: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep:\*

6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

7: /cgn2\_6/ptodata/1/1aa/6D.COMB.pep:\*

8: /cgn2\_6/ptodata/1/1aa/6E.COMB.pep:\*

9: /cgn2\_6/ptodata/1/1aa/6F.COMB.pep:\*

10: /cgn2\_6/ptodata/1/1aa/6G.COMB.pep:\*

11: /cgn2\_6/ptodata/1/1aa/6H.COMB.pep:\*

12: /cgn2\_6/ptodata/1/1aa/6I.COMB.pep:\*

13: /cgn2\_6/ptodata/1/1aa/6J.COMB.pep:\*

14: /cgn2\_6/ptodata/1/1aa/6K.COMB.pep:\*

15: /cgn2\_6/ptodata/1/1aa/6L.COMB.pep:\*

16: /cgn2\_6/ptodata/1/1aa/6M.COMB.pep:\*

17: /cgn2\_6/ptodata/1/1aa/6N.COMB.pep:\*

18: /cgn2\_6/ptodata/1/1aa/6O.COMB.pep:\*

19: /cgn2\_6/ptodata/1/1aa/6P.COMB.pep:\*

20: /cgn2\_6/ptodata/1/1aa/6Q.COMB.pep:\*

21: /cgn2\_6/ptodata/1/1aa/6R.COMB.pep:\*

22: /cgn2\_6/ptodata/1/1aa/6S.COMB.pep:\*

23: /cgn2\_6/ptodata/1/1aa/6T.COMB.pep:\*

24: /cgn2\_6/ptodata/1/1aa/6U.COMB.pep:\*

25: /cgn2\_6/ptodata/1/1aa/6V.COMB.pep:\*

26: /cgn2\_6/ptodata/1/1aa/6W.COMB.pep:\*

27: /cgn2\_6/ptodata/1/1aa/6X.COMB.pep:\*

28: /cgn2\_6/ptodata/1/1aa/6Y.COMB.pep:\*

29: /cgn2\_6/ptodata/1/1aa/6Z.COMB.pep:\*

30: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*

31: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*

32: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep:\*

33: /cgn2\_6/ptodata/1/1aa/6D.COMB.pep:\*

34: /cgn2\_6/ptodata/1/1aa/6E.COMB.pep:\*

35: /cgn2\_6/ptodata/1/1aa/6F.COMB.pep:\*

36: /cgn2\_6/ptodata/1/1aa/6G.COMB.pep:\*

37: /cgn2\_6/ptodata/1/1aa/6H.COMB.pep:\*

38: /cgn2\_6/ptodata/1/1aa/6I.COMB.pep:\*

39: /cgn2\_6/ptodata/1/1aa/6J.COMB.pep:\*

40: /cgn2\_6/ptodata/1/1aa/6K.COMB.pep:\*

41: /cgn2\_6/ptodata/1/1aa/6L.COMB.pep:\*

42: /cgn2\_6/ptodata/1/1aa/6M.COMB.pep:\*

43: /cgn2\_6/ptodata/1/1aa/6N.COMB.pep:\*

44: /cgn2\_6/ptodata/1/1aa/6O.COMB.pep:\*

45: /cgn2\_6/ptodata/1/1aa/6P.COMB.pep:\*

46: /cgn2\_6/ptodata/1/1aa/6Q.COMB.pep:\*

47: /cgn2\_6/ptodata/1/1aa/6R.COMB.pep:\*

48: /cgn2\_6/ptodata/1/1aa/6S.COMB.pep:\*

49: /cgn2\_6/ptodata/1/1aa/6T.COMB.pep:\*

50: /cgn2\_6/ptodata/1/1aa/6U.COMB.pep:\*

51: /cgn2\_6/ptodata/1/1aa/6V.COMB.pep:\*

52: /cgn2\_6/ptodata/1/1aa/6W.COMB.pep:\*

53: /cgn2\_6/ptodata/1/1aa/6X.COMB.pep:\*

54: /cgn2\_6/ptodata/1/1aa/6Y.COMB.pep:\*

55: /cgn2\_6/ptodata/1/1aa/6Z.COMB.pep:\*

56: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*

57: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*

58: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep:\*

59: /cgn2\_6/ptodata/1/1aa/6D.COMB.pep:\*

60: /cgn2\_6/ptodata/1/1aa/6E.COMB.pep:\*

61: /cgn2\_6/ptodata/1/1aa/6F.COMB.pep:\*

62: /cgn2\_6/ptodata/1/1aa/6G.COMB.pep:\*

63: /cgn2\_6/ptodata/1/1aa/6H.COMB.pep:\*

64: /cgn2\_6/ptodata/1/1aa/6I.COMB.pep:\*

65: /cgn2\_6/ptodata/1/1aa/6J.COMB.pep:\*

66: /cgn2\_6/ptodata/1/1aa/6K.COMB.pep:\*

67: /cgn2\_6/ptodata/1/1aa/6L.COMB.pep:\*

68: /cgn2\_6/ptodata/1/1aa/6M.COMB.pep:\*

69: /cgn2\_6/ptodata/1/1aa/6N.COMB.pep:\*

70: /cgn2\_6/ptodata/1/1aa/6O.COMB.pep:\*

71: /cgn2\_6/ptodata/1/1aa/6P.COMB.pep:\*

72: /cgn2\_6/ptodata/1/1aa/6Q.COMB.pep:\*

73: /cgn2\_6/ptodata/1/1aa/6R.COMB.pep:\*

74: /cgn2\_6/ptodata/1/1aa/6S.COMB.pep:\*

75: /cgn2\_6/ptodata/1/1aa/6T.COMB.pep:\*

76: /cgn2\_6/ptodata/1/1aa/6U.COMB.pep:\*

77: /cgn2\_6/ptodata/1/1aa/6V.COMB.pep:\*

78: /cgn2\_6/ptodata/1/1aa/6W.COMB.pep:\*

79: /cgn2\_6/ptodata/1/1aa/6X.COMB.pep:\*

80: /cgn2\_6/ptodata/1/1aa/6Y.COMB.pep:\*

81: /cgn2\_6/ptodata/1/1aa/6Z.COMB.pep:\*

82: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*

83: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*

84: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep:\*

85: /cgn2\_6/ptodata/1/1aa/6D.COMB.pep:\*

86: /cgn2\_6/ptodata/1/1aa/6E.COMB.pep:\*

87: /cgn2\_6/ptodata/1/1aa/6F.COMB.pep:\*

88: /cgn2\_6/ptodata/1/1aa/6G.COMB.pep:\*

89: /cgn2\_6/ptodata/1/1aa/6H.COMB.pep:\*

90: /cgn2\_6/ptodata/1/1aa/6I.COMB.pep:\*

91: /cgn2\_6/ptodata/1/1aa/6J.COMB.pep:\*

92: /cgn2\_6/ptodata/1/1aa/6K.COMB.pep:\*

93: /cgn2\_6/ptodata/1/1aa/6L.COMB.pep:\*

94: /cgn2\_6/ptodata/1/1aa/6M.COMB.pep:\*

95: /cgn2\_6/ptodata/1/1aa/6N.COMB.pep:\*

96: /cgn2\_6/ptodata/1/1aa/6O.COMB.pep:\*

97: /cgn2\_6/ptodata/1/1aa/6P.COMB.pep:\*

98: /cgn2\_6/ptodata/1/1aa/6Q.COMB.pep:\*

99: /cgn2\_6/ptodata/1/1aa/6R.COMB.pep:\*

100: /cgn2\_6/ptodata/1/1aa/6S.COMB.pep:\*

101: /cgn2\_6/ptodata/1/1aa/6T.COMB.pep:\*

102: /cgn2\_6/ptodata/1/1aa/6U.COMB.pep:\*

103: /cgn2\_6/ptodata/1/1aa/6V.COMB.pep:\*

104: /cgn2\_6/ptodata/1/1aa/6W.COMB.pep:\*

105: /cgn2\_6/ptodata/1/1aa/6X.COMB.pep:\*

106: /cgn2\_6/ptodata/1/1aa/6Y.COMB.pep:\*

107: /cgn2\_6/ptodata/1/1aa/6Z.COMB.pep:\*

108: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*

109: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*

110: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep:\*

111: /cgn2\_6/ptodata/1/1aa/6D.COMB.pep:\*

112: /cgn2\_6/ptodata/1/1aa/6E.COMB.pep:\*

113: /cgn2\_6/ptodata/1/1aa/6F.COMB.pep:\*

114: /cgn2\_6/ptodata/1/1aa/6G.COMB.pep:\*

115: /cgn2\_6/ptodata/1/1aa/6H.COMB.pep:\*

116: /cgn2\_6/ptodata/1/1aa/6I.COMB.pep:\*

117: /cgn2\_6/ptodata/1/1aa/6J.COMB.pep:\*

118: /cgn2\_6/ptodata/1/1aa/6K.COMB.pep:\*

119: /cgn2\_6/ptodata/1/1aa/6L.COMB.pep:\*

120: /cgn2\_6/ptodata/1/1aa/6M.COMB.pep:\*

121: /cgn2\_6/ptodata/1/1aa/6N.COMB.pep:\*

122: /cgn2\_6/ptodata/1/1aa/6O.COMB.pep:\*

123: /cgn2\_6/ptodata/1/1aa/6P.COMB.pep:\*

124: /cgn2\_6/ptodata/1/1aa/6Q.COMB.pep:\*

125: /cgn2\_6/ptodata/1/1aa/6R.COMB.pep:\*

126: /cgn2\_6/ptodata/1/1aa/6S.COMB.pep:\*

127: /cgn2\_6/ptodata/1/1aa/6T.COMB.pep:\*

128: /cgn2\_6/ptodata/1/1aa/6U.COMB.pep:\*

129: /cgn2\_6/ptodata/1/1aa/6V.COMB.pep:\*

130: /cgn2\_6/ptodata/1/1aa/6W.COMB.pep:\*

131: /cgn2\_6/ptodata/1/1aa/6X.COMB.pep:\*

132: /cgn2\_6/ptodata/1/1aa/6Y.COMB.pep:\*

133: /cgn2\_6/ptodata/1/1aa/6Z.COMB.pep:\*

134: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*

135: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*

136: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep:\*

137: /cgn2\_6/ptodata/1/1aa/6D.COMB.pep:\*

138: /cgn2\_6/ptodata/1/1aa/6E.COMB.pep:\*

139: /cgn2\_6/ptodata/1/1aa/6F.COMB.pep:\*

140: /cgn2\_6/ptodata/1/1aa/6G.COMB.pep:\*

141: /cgn2\_6/ptodata/1/1aa/6H.COMB.pep:\*

142: /cgn2\_6/ptodata/1/1aa/6I.COMB.pep:\*

143: /cgn2\_6/ptodata/1/1aa/6J.COMB.pep:\*

144: /cgn2\_6/ptodata/1/1aa/6K.COMB.pep:\*

145: /cgn2\_6/ptodata/1/1aa/6L.COMB.pep:\*

146: /cgn2\_6/ptodata/1/1aa/6M.COMB.pep:\*

147: /cgn2\_6/ptodata/1/1aa/6N.COMB.pep:\*

148: /cgn2\_6/ptodata/1/1aa/6O.COMB.pep:\*

149: /cgn2\_6/ptodata/1/1aa/6P.COMB.pep:\*

150: /cgn2\_6/ptodata/1/1aa/6Q.COMB.pep:\*

151: /cgn2\_6/ptodata/1/1aa/6R.COMB.pep:\*

152: /cgn2\_6/ptodata/1/1aa/6S.COMB.pep:\*

153: /cgn2\_6/ptodata/1/1aa/6T.COMB.pep:\*

154: /cgn2\_6/ptodata/1/1aa/6U.COMB.pep:\*

155: /cgn2\_6/ptodata/1/1aa/6V.COMB.pep:\*

156: /cgn2\_6/ptodata/1/1aa/6W.COMB.pep:\*

157: /cgn2\_6/ptodata/1/1aa/6X.COMB.pep:\*

158: /cgn2\_6/ptodata/1/1aa/6Y.COMB.pep:\*

159: /cgn2\_6/ptodata/1/1aa/6Z.COMB.pep:\*

160: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*

161: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*

162: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep:\*

163: /cgn2\_6/ptodata/1/1aa/6D.COMB.pep:\*

164: /cgn2\_6/ptodata/1/1aa/6E.COMB.pep:\*

165: /cgn2\_6/ptodata/1/1aa/6F.COMB.pep:\*

166: /cgn2\_6/ptodata/1/1aa/6G.COMB.pep:\*

167: /cgn2\_6/ptodata/1/1aa/6H.COMB.pep:\*

168: /cgn2\_6/ptodata/1/1aa/6I.COMB.pep:\*

169: /cgn2\_6/ptodata/1/1aa/6J.COMB.pep:\*

170: /cgn2\_6/ptodata/1/1aa/6K.COMB.pep:\*

171: /cgn2\_6/ptodata/1/1aa/6L.COMB.pep:\*

172: /cgn2\_6/ptodata/1/1aa/6M.COMB.pep:\*

173: /cgn2\_6/ptodata/1/1aa/6N.COMB.pep:\*

174: /cgn2\_6/ptodata/1/1aa/6O.COMB.pep:\*

175: /cgn2\_6/ptodata/1/1aa/6P.COMB.pep:\*

176: /cgn2\_6/ptodata/1/1aa/6Q.COMB.pep:\*

177: /cgn2\_6/ptodata/1/1aa/6R.COMB.pep:\*

178: /cgn2\_6/ptodata/1/1aa/6S.COMB.pep:\*

179: /cgn2\_6/ptodata/1/1aa/6T.COMB.pep:\*

180: /cgn2\_6/ptodata/1/1aa/6U.COMB.pep:\*

181: /cgn2\_6/ptodata/1/1aa/6V.COMB.pep:\*

182: /cgn2\_6/ptodata/1/1aa/6W.COMB.pep:\*

183: /cgn2\_6/ptodata/1/1aa/6X.COMB.pep:\*

184: /cgn2\_6/ptodata/1/1aa/6Y.COMB.pep:\*

185: /cgn2\_6/ptodata/1/1aa/6Z.COMB.pep:\*

186: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*

187: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*

188: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep:\*

189: /cgn2\_6/ptodata/1/1aa/6D.COMB.pep:\*

190: /cgn2\_6/ptodata/1/1aa/6E.COMB.pep:\*

191: /cgn2\_6/ptodata/1/1aa/6F.COMB.pep:\*

192: /cgn2\_6/ptodata/1/1aa/6G.COMB.pep:\*

193: /cgn2\_6/ptodata/1/1aa/6H.COMB.pep:\*

194: /cgn2\_6/ptodata/1/1aa/6I.COMB.pep:\*

195: /cgn2\_6/ptodata/1/1aa/6J.COMB.pep:\*

196: /cgn2\_6/ptodata/1/1aa/6K.COMB.pep:\*

197: /cgn2\_6/ptodata/1/1aa/6L.COMB.pep:\*

198: /cgn2\_6/ptodata/1/1aa/6M.COMB.pep:\*

199: /cgn2\_6/ptodata/1/1aa/6N.COMB.pep:\*

## ALIGNMENTS

```
RESULT 1
US-09-221-235-5
; Sequence 5, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,235
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
LENGTH: 455
TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-235-5

Query Match          100.0%; Score 455; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLGASFWQIKFDLQFFENCGGSGFSVYRAKWIISQDKEVAVKLLKIEKEAIIISVL 60
DB 1 MSLGASFWQIKFDLQFFENCGGSGFSVYRAKWIISQDKEVAVKLLKIEKEAIIISVL 60
QY 61 SHRNIIOFYGVILPEPNYGVITEYASLSGLYDINSNSEMDMDHIMTATDVAKGMHY 120
DB 61 SHRNIIOFYGVILPEPNYGVITEYASLSGLYDINSNSEMDMDHIMTATDVAKGMHY 120
QY 121 LHMEAPVKVIHRDLKSRNVVIADGVLTCDPGASRFHNHTTMSLVGTFFPMMAPEVIOS 180
DB 121 LHMEAPVKVIHRDLKSRNVVIADGVLTCDPGASRFHNHTTMSLVGTFFPMMAPEVIOS 180
QY 181 LPVSETCDTYSYGVILWEMLTREVPFKGLEGLQVAVLVVEKNERLTIPSSCPSPFAELIH 240
DB 181 LPVSETCDTYSYGVILWEMLTREVPFKGLEGLQVAVLVVEKNERLTIPSSCPSPFAELIH 240
QY 241 QCWEADAKKRPSPFKQIISLESMSNDTSLPDKCNSTFLHNKAERCEIATELERLKLERD 300
DB 241 QCWEADAKKRPSPFKQIISLESMSNDTSLPDKCNSTFLHNKAERCEIATELERLKLERD 300
QY 301 LSFKEQELKERERRLKMWQOKLTQOSNTPLPLAARMSEESYFESKTESNSAEMSCQI 360
DB 301 LSFKEQELKERERRLKMWQOKLTQOSNTPLPLAARMSEESYFESKTESNSAEMSCQI 360
QY 361 TATSNGEGHGNPISLQAMLMGFGDIFSNMKAQAVHSGQIMQAKONSSTTTSRRGK 420
DB 361 TATSNGEGHGNPISLQAMLMGFGDIFSNMKAQAVHSGQIMQAKONSSTTTSRRGK 420
QY 421 KVNVALGFSDFDLSEGDGDDDDDDGEEEDNDMDNSE 455
DB 421 KVNVALGFSDFDLSEGDGDDDDDDGEEEDNDMDNSE 455

RESULT 2
US-09-221-928-5
; Sequence 5, Application US/09221928
; Patent No. 6121030
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,928
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
```

```
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
LENGTH: 455
TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-928-5

Query Match          100.0%; Score 455; DB 3; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLGASFWQIKFDLQFFENCGGSGFSVYRAKWIISQDKEVAVKLLKIEKEAIIISVL 60
DB 1 MSLGASFWQIKFDLQFFENCGGSGFSVYRAKWIISQDKEVAVKLLKIEKEAIIISVL 60
QY 61 SHRNIIOFYGVILPEPNYGVITEYASLSGLYDINSNSEMDMDHIMTATDVAKGMHY 120
DB 61 SHRNIIOFYGVILPEPNYGVITEYASLSGLYDINSNSEMDMDHIMTATDVAKGMHY 120
QY 121 LHMEAPVKVIHRDLKSRNVVIADGVLTCDPGASRFHNHTTMSLVGTFFPMMAPEVIOS 180
DB 121 LHMEAPVKVIHRDLKSRNVVIADGVLTCDPGASRFHNHTTMSLVGTFFPMMAPEVIOS 180
QY 181 LPVSETCDTYSYGVILWEMLTREVPFKGLEGLQVAVLVVEKNERLTIPSSCPSPFAELIH 240
DB 181 LPVSETCDTYSYGVILWEMLTREVPFKGLEGLQVAVLVVEKNERLTIPSSCPSPFAELIH 240
QY 241 QCWEADAKKRPSPFKQIISLESMSNDTSLPDKCNSTFLHNKAERCEIATELERLKLERD 300
DB 241 QCWEADAKKRPSPFKQIISLESMSNDTSLPDKCNSTFLHNKAERCEIATELERLKLERD 300
QY 301 LSFKEQELKERERRLKMWQOKLTQOSNTPLPLAARMSEESYFESKTESNSAEMSCQI 360
DB 301 LSFKEQELKERERRLKMWQOKLTQOSNTPLPLAARMSEESYFESKTESNSAEMSCQI 360
QY 361 TATSNGEGHGNPISLQAMLMGFGDIFSNMKAQAVHSGQIMQAKONSSTTTSRRGK 420
DB 361 TATSNGEGHGNPISLQAMLMGFGDIFSNMKAQAVHSGQIMQAKONSSTTTSRRGK 420
QY 421 KVNVALGFSDFDLSEGDGDDDDDDGEEEDNDMDNSE 455
DB 421 KVNVALGFSDFDLSEGDGDDDDDDGEEEDNDMDNSE 455

RESULT 3
US-09-221-527-5
; Sequence 5, Application US/09221527
; Patent No. 6146832
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,527
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
LENGTH: 455
TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-527-5

Query Match          100.0%; Score 455; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLGASFWQIKFDLQFFENCGGSGFSVYRAKWIISQDKEVAVKLLKIEKEAIIISVL 60
DB 1 MSLGASFWQIKFDLQFFENCGGSGFSVYRAKWIISQDKEVAVKLLKIEKEAIIISVL 60
```

Qy 61 SHRNIIQFYGVILEPPNYGIVTEYASIGSLYDIYINSNRSEMDMDHIMTATDVAKGMY 120  
 Db 61 SHRNIIQFYGVILEPPNYGIVTEYASIGSLYDIYINSNRSEMDMDHIMTATDVAKGMY 120  
 Qy 121 LHMEAPVKYIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPEVIO 180  
 Db 121 LHMEAPVKYIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPEVIO 180  
 Qy 181 LPVSETCDITYSYGVVLEWMLTREVPFKGLEGLQVAMLVVEKNERLTIIPSSCPSPFAELH 240  
 Db 181 LPVSETCDITYSYGVVLEWMLTREVPFKGLEGLQVAMLVVEKNERLTIIPSSCPSPFAELH 240  
 Qy 241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNSEFLHNKAEMRCIEATLERLKKLERD 300  
 Db 241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNSEFLHNKAEMRCIEATLERLKKLERD 300  
 Qy 301 LSPKEQELKERERRLKMWEOKLTBOSNTPLLLPLAARMSSESYFESKTEESNSAEMSCOI 360  
 Db 301 LSPKEQELKERERRLKMWEOKLTBOSNTPLLLPLAARMSSESYFESKTEESNSAEMSCOI 360  
 Qy 361 TATSNGEGHGNPISLOAMLMGFQDIFSMNKAGAVHSGMOINMOAKONSSKTTSKRGK 420  
 Db 361 TATSNGEGHGNPISLOAMLMGFQDIFSMNKAGAVHSGMOINMOAKONSSKTTSKRGK 420  
 Qy 421 KVNMAFGSDFDLSEGDGDDDDGDEEDNDMDNSE 455  
 Db 421 KVNMAFGSDFDLSEGDGDDDDGDEEDNDMDNSE 455

## RESULT 4

US-09-221-236-5  
 ; Sequence 5, Application US/09221236  
 ; Patent No. 6146841  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Acton, Susan  
 ; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
 ; FILE REFERENCE: NMT-050  
 ; CURRENT APPLICATION NUMBER: US/09/221,236  
 ; EARLIER FILING DATE: 1998-12-28  
 ; EARLIER APPLICATION NUMBER: 09/163,115  
 ; EARLIER FILING DATE: 1998-09-29  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 5  
 ; LENGTH: 455  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 -09-221-236-5

Query Match Best Local Similarity 100.0%; Score 455; DB 4; Length 455;  
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSSIGASFVQIKFDDLOFPENCGGSGFSYVYRAKWSODKEVAVKLLIKEAEILLSVL 60  
 Db 1 MSSIGASFVQIKFDDLOFPENCGGSGFSYVYRAKWSODKEVAVKLLIKEAEILLSVL 60  
 Qy 61 SHRNIIQFYGVILEPPNYGIVTEYASIGSLYDIYINSNRSEMDMDHIMTATDVAKGMY 120  
 Db 61 SHRNIIQFYGVILEPPNYGIVTEYASIGSLYDIYINSNRSEMDMDHIMTATDVAKGMY 120  
 Qy 121 LHMEAPVKYIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPEVIO 180  
 Db 121 LHMEAPVKYIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPEVIO 180  
 Qy 181 LPVSETCDITYSYGVVLEWMLTREVPFKGLEGLQVAMLVVEKNERLTIIPSSCPSPFAELH 240  
 Db 181 LPVSETCDITYSYGVVLEWMLTREVPFKGLEGLQVAMLVVEKNERLTIIPSSCPSPFAELH 240  
 Qy 241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNSEFLHNKAEMRCIEATLERLKKLERD 300  
 Db 241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNSEFLHNKAEMRCIEATLERLKKLERD 300

Qy 301 LSPKEQELKERERRLKMWEOKLTBOSNTPLLLPLAARMSSESYFESKTEESNSAEMSCOI 360  
 Db 301 LSPKEQELKERERRLKMWEOKLTBOSNTPLLLPLAARMSSESYFESKTEESNSAEMSCOI 360  
 Qy 361 TATSNGEGHGNPISLOAMLMGFQDIFSMNKAGAVHSGMOINMOAKONSSKTTSKRGK 420  
 Db 361 TATSNGEGHGNPISLOAMLMGFQDIFSMNKAGAVHSGMOINMOAKONSSKTTSKRGK 420  
 Qy 421 KVNMAFGSDFDLSEGDGDDDDGDEEDNDMDNSE 455  
 Db 421 KVNMAFGSDFDLSEGDGDDDDGDEEDNDMDNSE 455

## RESULT 5

US-09-221-416-5  
 ; Sequence 5, Application US/09221416  
 ; Patent No. 6153417  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Acton, Susan  
 ; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
 ; FILE REFERENCE: NMT-050  
 ; CURRENT APPLICATION NUMBER: US/09/221,416  
 ; EARLIER FILING DATE: 1998-12-28  
 ; EARLIER APPLICATION NUMBER: 09/163,115  
 ; EARLIER FILING DATE: 1998-09-29  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 5  
 ; LENGTH: 455  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-221-416-5

Query Match Best Local Similarity 100.0%; Score 455; DB 4; Length 455;  
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSSIGASFVQIKFDDLOFPENCGGSGFSYVYRAKWSODKEVAVKLLIKEAEILLSVL 60  
 Db 1 MSSIGASFVQIKFDDLOFPENCGGSGFSYVYRAKWSODKEVAVKLLIKEAEILLSVL 60  
 Qy 61 SHRNIIQFYGVILEPPNYGIVTEYASIGSLYDIYINSNRSEMDMDHIMTATDVAKGMY 120  
 Db 61 SHRNIIQFYGVILEPPNYGIVTEYASIGSLYDIYINSNRSEMDMDHIMTATDVAKGMY 120  
 Qy 121 LHMEAPVKYIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPEVIO 180  
 Db 121 LHMEAPVKYIHRDLKSRNVVIAADGVLCIDFGASRFHNHTTHMSLVGTFPMMAPEVIO 180  
 Qy 181 LPVSETCDITYSYGVVLEWMLTREVPFKGLEGLQVAMLVVEKNERLTIIPSSCPSPFAELH 240  
 Db 181 LPVSETCDITYSYGVVLEWMLTREVPFKGLEGLQVAMLVVEKNERLTIIPSSCPSPFAELH 240  
 Qy 241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNSEFLHNKAEMRCIEATLERLKKLERD 300  
 Db 241 QCWEADAKKRPSPFKQIISILESMSNDTSLPDKCNSEFLHNKAEMRCIEATLERLKKLERD 300  
 Qy 301 LSPKEQELKERERRLKMWEOKLTBOSNTPLLLPLAARMSSESYFESKTEESNSAEMSCOI 360  
 Db 301 LSPKEQELKERERRLKMWEOKLTBOSNTPLLLPLAARMSSESYFESKTEESNSAEMSCOI 360  
 Qy 361 TATSNGEGHGNPISLOAMLMGFQDIFSMNKAGAVHSGMOINMOAKONSSKTTSKRGK 420  
 Db 361 TATSNGEGHGNPISLOAMLMGFQDIFSMNKAGAVHSGMOINMOAKONSSKTTSKRGK 420  
 Qy 421 KVNMAFGSDFDLSEGDGDDDDGDEEDNDMDNSE 455  
 Db 421 KVNMAFGSDFDLSEGDGDDDDGDEEDNDMDNSE 455

## RESULT 6

US-09-221-245-5

```

; Sequence 5, Application US/09221245
; Patent No. 6180358
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,245
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: US 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PR
; ORGANISM: Homo sapiens
US-09-221-245-5

```

```

Query Match          100.0%; Score 455; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MSSLGASFYQIKFDLLOPFENCGGSGFSYVRAKWISODKEVAVKLKIKEKAEIISVL 60
DB 1 MSSLGASFYQIKFDLLOPFENCGGSGFSYVRAKWISODKEVAVKLKIKEKAEIISVL 60
QY 61 SHRNIIOFYGVILPEPNYGIITEVASLSGLYDINSNSSEEMDMHMTWATDVAKGMHY 120
DB 61 SHRNIIOFYGVILPEPNYGIITEVASLSGLYDINSNSSEEMDMHMTWATDVAKGMHY 120
QY 121 LHMEAPVKYIHRDLKSRNVVIAADGVLCIDFGASRFNNHTTHMSLVGTFPMMADEVIOS 180
DB 121 LHMEAPVKYIHRDLKSRNVVIAADGVLCIDFGASRFNNHTTHMSLVGTFPMMADEVIOS 180
QY 121 LHMEAPVKYIHRDLKSRNVVIAADGVLCIDFGASRFNNHTTHMSLVGTFPMMADEVIOS 180
DB 121 LHMEAPVKYIHRDLKSRNVVIAADGVLCIDFGASRFNNHTTHMSLVGTFPMMADEVIOS 180
QY 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEQVAMLVKNERLTIPSSCRSPFAELH 240
DB 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEQVAMLVKNERLTIPSSCRSPFAELH 240
QY 241 QCWEADAKKRPSPKQIISLESMSNDTSLPDKNSFLHKAEMRCEIATELRLKLERD 300
DB 241 QCWEADAKKRPSPKQIISLESMSNDTSLPDKNSFLHKAEMRCEIATELRLKLERD 300
QY 301 LSPKEQELKERERRLKMEQKLTQOSNTPLLPLAARMEESYFESKTESNSAEMSCOI 360
DB 301 LSPKEQELKERERRLKMEQKLTQOSNTPLLPLAARMEESYFESKTESNSAEMSCOI 360
QY 361 TATSNGEGHGMNPSLOAMLMGFGDIFSMNKAGAVHSGMOINMOAKONSSKTTSKRGK 420
DB 361 TATSNGEGHGMNPSLOAMLMGFGDIFSMNKAGAVHSGMOINMOAKONSSKTTSKRGK 420
QY 421 KVNMAIGFSDFDLSEGDDEDDDDGEEEDNDMDNSE 455
DB 421 KVNMAIGFSDFDLSEGDDEDDDDGEEEDNDMDNSE 455

```

```

RESULT 7
US-09-163-115-5
; Sequence 5, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PR
; ORGANISM: Homo sapiens
US-09-163-115-5

```

```

Query Match          100.0%; Score 455; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MSSLGASFYQIKFDLLOPFENCGGSGFSYVRAKWISODKEVAVKLKIKEKAEIISVL 60
DB 1 MSSLGASFYQIKFDLLOPFENCGGSGFSYVRAKWISODKEVAVKLKIKEKAEIISVL 60
QY 61 SHRNIIOFYGVILPEPNYGIITEVASLSGLYDINSNSSEEMDMHMTWATDVAKGMHY 120
DB 61 SHRNIIOFYGVILPEPNYGIITEVASLSGLYDINSNSSEEMDMHMTWATDVAKGMHY 120
QY 121 LHMEAPVKYIHRDLKSRNVVIAADGVLCIDFGASRFNNHTTHMSLVGTFPMMADEVIOS 180
DB 121 LHMEAPVKYIHRDLKSRNVVIAADGVLCIDFGASRFNNHTTHMSLVGTFPMMADEVIOS 180
QY 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEQVAMLVKNERLTIPSSCRSPFAELH 240
DB 181 LPVSETCDTYSYGVVLMEMLTREVPFKGLEQVAMLVKNERLTIPSSCRSPFAELH 240
QY 241 QCWEADAKKRPSPKQIISLESMSNDTSLPDKNSFLHKAEMRCEIATELRLKLERD 300
DB 241 QCWEADAKKRPSPKQIISLESMSNDTSLPDKNSFLHKAEMRCEIATELRLKLERD 300
QY 301 LSPKEQELKERERRLKMEQKLTQOSNTPLLPLAARMEESYFESKTESNSAEMSCOI 360
DB 301 LSPKEQELKERERRLKMEQKLTQOSNTPLLPLAARMEESYFESKTESNSAEMSCOI 360
QY 361 TATSNGEGHGMNPSLOAMLMGFGDIFSMNKAGAVHSGMOINMOAKONSSKTTSKRGK 420
DB 361 TATSNGEGHGMNPSLOAMLMGFGDIFSMNKAGAVHSGMOINMOAKONSSKTTSKRGK 420
QY 421 KVNMAIGFSDFDLSEGDDEDDDDGEEEDNDMDNSE 455
DB 421 KVNMAIGFSDFDLSEGDDEDDDDGEEEDNDMDNSE 455

```

```

RESULT 8
US-09-221-528-5
; Sequence 5, Application US/09221528
; Patent No. 6190874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,528
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PR
; ORGANISM: Homo sapiens
US-09-221-528-5

```

```

Query Match          100.0%; Score 455; DB 4; Length 455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MSSLGASFYQIKFDLLOPFENCGGSGFSYVRAKWISODKEVAVKLKIKEKAEIISVL 60
DB 1 MSSLGASFYQIKFDLLOPFENCGGSGFSYVRAKWISODKEVAVKLKIKEKAEIISVL 60
QY 61 SHRNIIOFYGVILPEPNYGIITEVASLSGLYDINSNSSEEMDMHMTWATDVAKGMHY 120
DB 61 SHRNIIOFYGVILPEPNYGIITEVASLSGLYDINSNSSEEMDMHMTWATDVAKGMHY 120
QY 121 LHMEAPVKYIHRDLKSRNVVIAADGVLCIDFGASRFNNHTTHMSLVGTFPMMADEVIOS 180
DB 121 LHMEAPVKYIHRDLKSRNVVIAADGVLCIDFGASRFNNHTTHMSLVGTFPMMADEVIOS 180

```



QY 181 LPVSETCTYSGVIVLWMLTREVPFKGLGLOVAMLVKNERLTTPSSCPSPFAELH 240  
 DB 181 LPVSETCTYSGVIVLWMLTREVPFKGLGLOVAMLVKNERLTTPSSCPSPFAELH 240  
 QY 241 QCEWADAKKRPSPFOIISILESMSNDTSLPDKNSFLHNKAEMWCEIATLERLKLKLERD 300  
 DB 241 QCEWADAKKRPSPFOIISILESMSNDTSLPDKNSFLHNKAEMWCEIATLERLKLKLERD 300  
 QY 301 LSPFEQELKERERRLKMWEOKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 360  
 DB 301 LSPFEQELKERERRLKMWEOKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 360  
 QY 361 TATSNGEGHGNPBLQAMLMGFGDIFSNKAGAVHSGMOINMOAKONSKTTSKRKRG 420  
 DB 361 TATSNGEGHGNPBLQAMLMGFGDIFSNKAGAVHSGMOINMOAKONSKTTSKRKRG 420  
 QY 421 KVNMAIGFSDPDLSEGGDDDDDDGEEEDNDMDNSE 455  
 DB 421 KVNMAIGFSDPDLSEGGDDDDDDGEEEDNDMDNSE 455

## RESULT 9

US-09-593-553-5  
 ; Sequence 5, Application US/09593553  
 ; Patent No. 6200770  
 ; GENERAL INFORMATION:

APPLICANT: Acton, Susan  
 TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
 FILE REFERENCE: MNI-050  
 CURRENT APPLICATION NUMBER: US/09/593,553  
 CURRENT FILING DATE: 2000-06-14  
 PRIOR APPLICATION NUMBER: 09/163,115  
 PRIOR FILING DATE: 1998-09-28  
 NUMBER OF SEQ ID NOS: 15  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 5  
 LENGTH: 455  
 TYPE: PRF  
 ORGANISM: Homo sapiens  
 US-09-593-553-5

Query Match 100.0%; Score 455; DB 4; Length 455;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLSGAFVQIKFDLQFFENCGGSGFSVYRAKWIISQDEKVAVKLLKIEKEAELISVL 60  
 DB 1 MSLSGAFVQIKFDLQFFENCGGSGFSVYRAKWIISQDEKVAVKLLKIEKEAELISVL 60  
 QY 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYNINSRSEEMDMHMTWATDVAKGMHY 120  
 DB 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYNINSRSEEMDMHMTWATDVAKGMHY 120  
 QY 121 LHMEAPVKVIRHDLKSRNVVIADGVLIKIDFGASRPHNHTTHMSLVGTFPMMAPEVIQS 180  
 DB 121 LHMEAPVKVIRHDLKSRNVVIADGVLIKIDFGASRPHNHTTHMSLVGTFPMMAPEVIQS 180  
 QY 181 LPVSETCTYSGVIVLWMLTREVPFKGLGLOVAMLVKNERLTTPSSCPSPFAELH 240  
 DB 181 LPVSETCTYSGVIVLWMLTREVPFKGLGLOVAMLVKNERLTTPSSCPSPFAELH 240  
 QY 241 QCEWADAKKRPSPFOIISILESMSNDTSLPDKNSFLHNKAEMWCEIATLERLKLKLERD 300  
 DB 241 QCEWADAKKRPSPFOIISILESMSNDTSLPDKNSFLHNKAEMWCEIATLERLKLKLERD 300  
 QY 301 LSPFEQELKERERRLKMWEOKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 360  
 DB 301 LSPFEQELKERERRLKMWEOKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 360  
 QY 361 TATSNGEGHGNPBLQAMLMGFGDIFSNKAGAVHSGMOINMOAKONSKTTSKRKRG 420  
 DB 361 TATSNGEGHGNPBLQAMLMGFGDIFSNKAGAVHSGMOINMOAKONSKTTSKRKRG 420

QY 421 KVNMAIGFSDPDLSEGGDDDDDDGEEEDNDMDNSE 455  
 DB 421 KVNMAIGFSDPDLSEGGDDDDDDGEEEDNDMDNSE 455

## RESULT 10

US-09-221-237-5  
 ; Sequence 5, Application US/09221237  
 ; Patent No. 6214597  
 ; GENERAL INFORMATION:

APPLICANT: Acton, Susan  
 TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
 FILE REFERENCE: MNI-050  
 CURRENT APPLICATION NUMBER: US/09/221,237  
 EARLIER APPLICATION NUMBER: 09/163,115  
 EARLIER FILING DATE: 1998-09-29  
 NUMBER OF SEQ ID NOS: 15  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 5  
 LENGTH: 455  
 TYPE: PRF  
 ORGANISM: Homo sapiens  
 US-09-221-237-5

Query Match 100.0%; Score 455; DB 4; Length 455;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLSGAFVQIKFDLQFFENCGGSGFSVYRAKWIISQDEKVAVKLLKIEKEAELISVL 60  
 DB 1 MSLSGAFVQIKFDLQFFENCGGSGFSVYRAKWIISQDEKVAVKLLKIEKEAELISVL 60  
 QY 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYNINSRSEEMDMHMTWATDVAKGMHY 120  
 DB 61 SHRNIIQFYGVILEPPNYGIVTEYASLSGLYDYNINSRSEEMDMHMTWATDVAKGMHY 120  
 QY 121 LHMEAPVKVIRHDLKSRNVVIADGVLIKIDFGASRPHNHTTHMSLVGTFPMMAPEVIQS 180  
 DB 121 LHMEAPVKVIRHDLKSRNVVIADGVLIKIDFGASRPHNHTTHMSLVGTFPMMAPEVIQS 180  
 QY 181 LPVSETCTYSGVIVLWMLTREVPFKGLGLOVAMLVKNERLTTPSSCPSPFAELH 240  
 DB 181 LPVSETCTYSGVIVLWMLTREVPFKGLGLOVAMLVKNERLTTPSSCPSPFAELH 240  
 QY 241 QCEWADAKKRPSPFOIISILESMSNDTSLPDKNSFLHNKAEMWCEIATLERLKLKLERD 300  
 DB 241 QCEWADAKKRPSPFOIISILESMSNDTSLPDKNSFLHNKAEMWCEIATLERLKLKLERD 300  
 QY 301 LSPFEQELKERERRLKMWEOKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 360  
 DB 301 LSPFEQELKERERRLKMWEOKLTQOSNTPLLLPLAARMESESYFESKTEESNSAEMSCQI 360  
 QY 361 TATSNGEGHGNPBLQAMLMGFGDIFSNKAGAVHSGMOINMOAKONSKTTSKRKRG 420  
 DB 361 TATSNGEGHGNPBLQAMLMGFGDIFSNKAGAVHSGMOINMOAKONSKTTSKRKRG 420  
 QY 421 KVNMAIGFSDPDLSEGGDDDDDDGEEEDNDMDNSE 455  
 DB 421 KVNMAIGFSDPDLSEGGDDDDDDGEEEDNDMDNSE 455

## RESULT 11

US-08-375-300-4  
 ; Sequence 4, Application US/08375300  
 ; Patent No. 5679566  
 ; GENERAL INFORMATION:

APPLICANT: Feng, He  
 TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN  
 TITLE OF INVENTION: THE ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION  
 NUMBER OF SEQUENCES: 6  
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street Suite 3100  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/375,300  
FILING DATE: 20-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fasse, J. P.  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 04020/046001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)542-5070  
TELEFAX: (617)542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 764 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-375-300-4

Query Match 2.2%; Score 10; DB 1; Length 764;  
Best Local Similarity 100.0%; Pred. No. 0.64;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGEE 446  
DB 577 DDDDDDDGEE 586

RESULT 12  
US-09-177-431-4  
Sequence 4, Application US/09177431  
Patent No. 6071700  
GENERAL INFORMATION:  
APPLICANT: He, Feng  
APPLICANT: Jacobson, Allan S.  
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE  
TITLE OF INVENTION: ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/177,431  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/955,472  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fasse, J. Peter  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 07917/050001  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070  
TELEFAX: 617/542-9806  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 764 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-177-431-4

Query Match 2.2%; Score 10; DB 3; Length 764;  
Best Local Similarity 100.0%; Pred. No. 0.64;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGEE 446  
DB 577 DDDDDDDGEE 586

RESULT 13  
PCT-US95-16930-4  
Sequence 4, Application PC/TUS9516930  
GENERAL INFORMATION:  
APPLICANT: UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL  
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE  
TITLE OF INVENTION: PRODUCTION IN THE ABSENCE OF  
TITLE OF INVENTION: NONSENSE-MEDIATED mRNA DECAY  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street Suite 3100  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/16930  
FILING DATE: 27-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/375,300  
FILING DATE: 20-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fasse, J. Peter  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 04020/046W01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)542-5070  
TELEFAX: (617)542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 764 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-16930-4

Query Match 2.2%; Score 10; DB 5; Length 764;  
Best Local Similarity 100.0%; Pred. No. 0.64;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDGEE 446  
DB 577 DDDDDDDGEE 586

RESULT 14  
US-08-375-300-2  
Sequence 2, Application US/08375300  
Patent No. 5679566  
GENERAL INFORMATION:  
APPLICANT: Feng, He  
APPLICANT: Jacobson, Allan S.  
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN  
TITLE OF INVENTION: THE ABSENCE OF NONSENSE-MEDIATED MRNA DECAY FUNCTION  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street Suite 3100  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/375,300  
FILING DATE: 20-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fasse, J. P.  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 04020/046001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)542-5070  
TELEFAX: (617)542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1089 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-375-300-2

Query Match 2.2%; Score 10; DB 1; Length 1089;  
Best Local Similarity 100.0%; Pred. No. 0.86;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

437 DDDDDGEE 446  
|||||  
902 DDDDDGEE 911

RESULT 15  
US-09-177-431-2  
Sequence 2, Application US/09177431  
Patent No. 6071700  
GENERAL INFORMATION:  
APPLICANT: He, Feng  
APPLICANT: Jacobson, Allan S.  
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE  
TITLE OF INVENTION: ABSENCE OF NONSENSE-MEDIATED MRNA DECAY FUNCTION  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows 95

SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/177,431  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/955,472  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fasse, J. Peter  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 07917/050001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-9806  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1089 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-177-431-2

Query Match 2.2%; Score 10; DB 3; Length 1089;  
Best Local Similarity 100.0%; Pred. No. 0.86;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDGEE 446  
|||||  
Db 902 DDDDDGEE 911

RESULT 16  
PCT-US95-16930-2  
Sequence 2, Application PC/TUS9516930  
GENERAL INFORMATION:  
APPLICANT: UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL  
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE  
TITLE OF INVENTION: PRODUCTION IN THE ABSENCE OF  
TITLE OF INVENTION: NONSENSE-MEDIATED MRNA DECAY  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street Suite 3100  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/16930  
FILING DATE: 27-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/375,300  
FILING DATE: 20-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Fasse, J. Peter  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 04020/046001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)542-5070  
TELEFAX: (617)542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1089 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant

TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-16930-2

Query Match 2.2%; Score 10; DB 5; Length 1089;  
Best Local Similarity 100.0%; Pred. No. 0.86;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDGEE 446  
|||||

Db 902 DDDDDGEE 911

## RESULT 17

US-08-158-735A-15  
Sequence 15, Application US/08158735A

Patent No. 6248554

GENERAL INFORMATION:

APPLICANT: COOK, JONATHAN S.

APPLICANT: CORREA, PAUL E.

APPLICANT: KOENIG, BETH B.

APPLICANT: ROSENBAUM, JAN S.

APPLICANT: TING, JERRY

TITLE OF INVENTION: DNA SEQUENCE CODING FOR A BMP RECEPTOR

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: THE PROCTER & GAMBLE COMPANY

STREET: 11810 EAST MIAMI RIVER ROAD

CITY: ROSS

STATE: OH

COUNTRY: USA

ZIP: 45061

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/158,735A

FILING DATE: 24-NOV-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: CORSTANTE, BRAH J.

REGISTRATION NUMBER: 34,804

REFERENCE/DOCKET NUMBER: 5088

TELECOMMUNICATION INFORMATION:

TELEPHONE: (513) 627-2858

TELEFAX: (513) 627-0260

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 303 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-158-735A-15

Query Match 2.0%; Score 9; DB 4; Length 303;

Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139  
|||||

Db 128 HRDLKSRNV 136

## RESULT 18

US-09-382-256-12  
Sequence 12, Application US/09382256A

Patent No. 6207814

GENERAL INFORMATION:

APPLICANT: MIYAZONO, Kohel

APPLICANT: TEN DIJKE, Peter

FRANZEN, Petia  
YAMASHITA, Hidetoshi

HELDIN, Carl-Henrik

TITLE OF INVENTION: ACTIVIN RECEPTOR LIKE KINASES, PROTEINS,

HAVING SERINE THREONINE KINASE DOMAINS,

AND THEIR USE

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fulbright & Jaworski L.L.P.

STREET: 666 Fifth Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10103

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.25 inch, 1.44mb

OPERATING SYSTEM: PC-DOS

SOFTWARE: wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/382,256A

FILING DATE: 24-Aug-1999

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB93/02367

FILING DATE: No. 6207814ember 17, 1993

APPLICATION NUMBER: GB 9224057.1

FILING DATE: No. 6207814ember 17, 1992

APPLICATION NUMBER: GB 9304677.9

FILING DATE: March 8, 1993

APPLICATION NUMBER: GB 9304680.3

FILING DATE: March 8, 1993

APPLICATION NUMBER: 9311047.6

FILING DATE: May 28, 1993

APPLICATION NUMBER: 9313763.6

FILING DATE: July 2, 1993

APPLICATION NUMBER: 9316099.2

FILING DATE: August 3, 1993

APPLICATION NUMBER: 321344.5

FILING DATE: October 15, 1993

ATTORNEY/AGENT INFORMATION:

NAME: No. 6207814man D. Hanson

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: LUD 5298.1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 318-3000

TELEFAX: (212) 752-5958

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 502 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-09-382-256-12

Query Match 2.0%; Score 9; DB 4; Length 502;

Best Local Similarity 100.0%; Pred. No. 3.6;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139  
|||||

Db 327 HRDLKSRNV 335

## RESULT 19

US-09-395-115-12  
Sequence 12, Application US/09395115

Patent No. 6271365

GENERAL INFORMATION:

APPLICANT: Miyazono, Kohel; DiJke, Peter Ten;

APPLICANT: Franzen, Petia; Yamashita, Hidetoshi; Heldin, Carl-Henrik

TITLE OF INVENTION: Activin Receptor-Like Kinase, Proteins

TITLE OF INVENTION: Having Serine Threonine Kinase Domains And Their Use  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/395,115  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/436,265  
FILING DATE: 30-October-1995  
APPLICATION NUMBER: PCT/GB93/02367  
FILING DATE: 17-No. 6271365ember-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9224057.1  
FILING DATE: 17-No. 6271365ember-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9304677.9  
FILING DATE: 8-March-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9304680.3  
FILING DATE: 8-March-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9311047.6  
FILING DATE: 28-May-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9313763.6  
FILING DATE: 2-July-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9136099.2  
FILING DATE: 3-August-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9321344.5  
FILING DATE: 15-October-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Kohlei, Vineet  
REGISTRATION NUMBER: 37,003  
REFERENCE/DOCKET NUMBER: LUD 5298  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 502 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-395-115-12

Query Match 2.0%; Score 9; DB 4; Length 502;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139  
|||||  
Db 327 HRDLKSRNV 335

RESULT 20  
US-08-436-265-12  
; Sequence 12, Application US/08436265  
; Patent No. 6316217  
; GENERAL INFORMATION:  
; APPLICANT: Miyazono, Kohlei, DiJke, Peter Ten;

APPLICANT: Franzen, Petra; Yamashita, Hidetoshi; Heldin, Carl-Henrik  
TITLE OF INVENTION: Activin Receptor-Like Kinase, Proteins  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/436,265  
FILING DATE: 30-October-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB93/02367  
FILING DATE: 17-No. 6316217ember-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9224057.1  
FILING DATE: 17-No. 6316217ember-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9304677.9  
FILING DATE: 8-March-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9304680.3  
FILING DATE: 8-March-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9311047.6  
FILING DATE: 28-May-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9313763.6  
FILING DATE: 2-July-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9136099.2  
FILING DATE: 3-August-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9321344.5  
FILING DATE: 15-October-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Kohlei, Vineet  
REGISTRATION NUMBER: 37,003  
REFERENCE/DOCKET NUMBER: LUD 5298  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 502 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-436-265-12

Query Match 2.0%; Score 9; DB 4; Length 502;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDLKSRNV 139  
|||||  
Db 327 HRDLKSRNV 335

RESULT 21  
US-09-679-187-12  
; Sequence 12, Application US/09679187  
; Patent No. 631621  
; GENERAL INFORMATION:  
; APPLICANT: Miyazono, Kohlei, DiJke, Peter Ten;

APPLICANT: Franzen, Petra; Yamashita, Hidetoshi; Heldin, Carl-Henrik  
TITLE OF INVENTION: Activin Receptor-Like Kinase, Proteins  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/679,187  
FILING DATE: 03-OCT-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/436,265  
FILING DATE: 30-October-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB93/02367  
FILING DATE: 17-No. 6331621ember-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9224057.1  
FILING DATE: 17-No. 6331621ember-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9304677.9  
FILING DATE: 8-March-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9304680.3  
FILING DATE: 8-March-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9311047.6  
FILING DATE: 28-May-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9313763.6  
FILING DATE: 2-July-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9136099.2  
FILING DATE: 3-August-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 9321344.5  
FILING DATE: 15-October-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Kohlei, Vineet  
REGISTRATION NUMBER: 37,003  
REFERENCE/DOCKET NUMBER: LUD 5298  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 502 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-679-187-12

Query Match 2.0%; Score 9; DB 4; Length 502;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDIKSRNV 139  
|||  
Db 327 HRDIKSRNV 335

RESULT 22  
US-08-149-105-16  
Sequence 16, Application US/08149105  
Patent No. 5538892

GENERAL INFORMATION:  
APPLICANT: Donahoe, Patricia K.  
APPLICANT: Gustafson, Michael  
APPLICANT: He, Wei W.  
APPLICANT: Wang, Xiao-Fan  
TITLE OF INVENTION: TGF- TYPE I RECEPTOR  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: Wordperfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/149,105  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/029,673  
FILING DATE: March 11, 1993  
APPLICATION NUMBER: 07/853,396  
FILING DATE: March 18, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/211001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 505  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-149-105-16

Query Match 2.0%; Score 9; DB 1; Length 505;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 HRDIKSRNV 139  
|||  
Db 330 HRDIKSRNV 338

RESULT 23  
US-08-317-847-16  
Sequence 16, Application US/08317847  
Patent No. 5547854  
GENERAL INFORMATION:  
APPLICANT: Donahoe, Patricia K.  
APPLICANT: Gustafson, Michael  
APPLICANT: He, Wei W.  
TITLE OF INVENTION: FOUR NOVEL RECEPTORS OF THE TGF-B  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM PS/2 Model 50Z or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/317,847  
FILING DATE:  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/029,673  
FILING DATE: March 11, 1993  
APPLICATION NUMBER: 07/853,396  
FILING DATE: March 18, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/127002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 505  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-317-847-16

Query Match 2.0%; Score 9; DB 1; Length 505;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 131 HDLKSRY 139  
DB 330 HDLKSRY 338

RESULT 24 \*  
US-08-446-855A-2  
Sequence 2, Application US/08446855A  
Patent No. 5849573  
GENERAL INFORMATION:  
APPLICANT: Stewart, Thomas S  
APPLICANT: Flores, Maria V  
APPLICANT: O'Sullivan, William J  
TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderhye PC  
STREET: 1100 NO. 5849573th Gledde Road, 8th Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446,855A  
FILING DATE: 06-Jul-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mitchard, Leonard C  
REGISTRATION NUMBER: 29,009  
REFERENCE/DOCKET NUMBER: 47-80  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:

LENGTH: 2391 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-446-855A-2

Query Match 2.0%; Score 9; DB 2; Length 2391;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDGE 445  
DB 1786 DDDDDDDGE 1794

RESULT 25  
US-09-150-741-2  
Sequence 2, Application US/09150741  
Patent No. 6183996  
GENERAL INFORMATION:  
APPLICANT: Stewart et al.  
TITLE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl Phosphate  
FILE REFERENCE:  
CURRENT APPLICATION NUMBER: US/09/150,741  
EARLIER FILING DATE: 1998-09-10  
EARLIER APPLICATION NUMBER: PL6380  
EARLIER FILING DATE: 1992-12-16  
EARLIER APPLICATION NUMBER: AU93/00617  
EARLIER FILING DATE: 1993-12-02  
EARLIER APPLICATION NUMBER: 08/446,855  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 2  
LENGTH: 2391  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: protein  
US-09-150-741-2

Query Match 2.0%; Score 9; DB 4; Length 2391;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 437 DDDDDDDGE 445  
DB 1786 DDDDDDDGE 1794

RESULT 26  
US-08-240-712-35  
Sequence 35, Application US/08240712  
Patent No. 5539907  
GENERAL INFORMATION:  
APPLICANT: ANDERSON, DAVID C.  
APPLICANT: MATHERS, ANTONY JAMES  
APPLICANT: STELLER, GARY L.  
TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Broadway and Nelmark  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/240.712  
FILING DATE: 09-MAY-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/09752  
FILING DATE: 13-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, IVER P  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: ANDERSON=6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
OTHER INFORMATION: /note= one or both of Gly  
OTHER INFORMATION: residues 2 and 3 can be absent; any or all  
OTHER INFORMATION: of Asp residues 5-33 can be absent; one  
OTHER INFORMATION: or both of Gly residues 35 and 36 can be  
OTHER INFORMATION: absent  
MOLECULE TYPE: peptide  
US-08-240-712-35  
Query Match 1.8%; Score 8; DB 1; Length 36;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 436 GDDDDDD 443  
DB 3 GDDDDDD 10  
RESULT 27  
US-08-443-890-35  
Sequence 35, Application US/08443890  
Patent No. 5739011  
GENERAL INFORMATION:  
APPLICANT: ANDERSON, DAVID C.  
APPLICANT: MATHEWS, ANTONY JAMES  
APPLICANT: STETLER, GARY L.  
TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC  
TITLE OF INVENTION: HEMOGLOBINS  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Broadway and Neimark  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/443,890  
FILING DATE: 31-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/240,712  
FILING DATE: 09-MAY-1994  
APPLICATION NUMBER: PCT/US92/09752

FILING DATE: 13-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: COOPER, IVER P  
REGISTRATION NUMBER: 28,005  
REFERENCE/DOCKET NUMBER: ANDERSON=6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
TELEX: 248633  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
OTHER INFORMATION: /note= one or both of Gly  
OTHER INFORMATION: residues 2 and 3 can be absent; any or all  
OTHER INFORMATION: of Asp residues 5-33 can be absent; one  
OTHER INFORMATION: or both of Gly residues 35 and 36 can be  
OTHER INFORMATION: absent  
MOLECULE TYPE: peptide  
US-08-443-890-35  
Query Match 1.8%; Score 8; DB 1; Length 36;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 436 GDDDDDD 443  
DB 3 GDDDDDD 10  
RESULT 28  
US-09-303-814-1  
Sequence 1, Application US/09303814  
Patent No. 6391848  
GENERAL INFORMATION:  
APPLICANT: de Lumen, Benito O.  
APPLICANT: Galvez, Alfredo F.  
TITLE OF INVENTION: Soybean Protein Nutraceuicals  
FILE REFERENCE: B99-083  
CURRENT APPLICATION NUMBER: US/09/303,814  
CURRENT FILING DATE: 1999-04-30  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 1  
LENGTH: 43  
TYPE: PRT  
ORGANISM: soybean  
US-09-303-814-1  
Query Match 1.8%; Score 8; DB 4; Length 43;  
Best Local Similarity 100.0%; Pred. No. 3.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 436 GDDDDDD 443  
DB 34 GDDDDDD 41  
RESULT 29  
US-08-618-911-4  
Sequence 4, Application US/08618911  
Patent No. 5850016  
GENERAL INFORMATION:  
APPLICANT: Jung, Rudolf  
APPLICANT: Hastings, Craig  
APPLICANT: Coughlan, Sean  
APPLICANT: Hu, David  
TITLE OF INVENTION: ALTERATION OF AMINO ACID COMPOSITIONS IN  
TITLE OF INVENTION: SEEDS  
NUMBER OF SEQUENCES: 13



;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.  
;; STREET: 700 Capital Square, 400 Locust Street  
;; CITY: Des Moines  
;; STATE: Iowa  
;; COUNTRY: USA  
;; ZIP: 50309  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/618,911  
;; FILING DATE: Concurrently herewith  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Simon, Soma  
;; REGISTRATION NUMBER: 37,444  
;; REFERENCE/DOCKET NUMBER: 365-US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (515) 248-4800  
;; TELEFAX: (515) 248-4844  
;;  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 158 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-618-911-4  
;;  
Query Match 1.8%; Score 8; DB 2; Length 158;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 436 GDDDDDD 443  
Db 55 GDDDDDD 62  
;;  
RESULT 30  
US-08-618-911-6  
;; Sequence 6, Application US/08618911  
;; Patent No. 5850016  
;; GENERAL INFORMATION:  
;; APPLICANT: Jung, Rudolf  
;; APPLICANT: Haefliger, Craig  
;; APPLICANT: Coughlan, Sean  
;; APPLICANT: Hu, David  
;; TITLE OF INVENTION: ALTERATION OF AMINO ACID COMPOSITIONS IN  
;; TITLE OF INVENTION: SEEDS  
;; NUMBER OF SEQUENCES: 13  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.  
;; STREET: 700 Capital Square, 400 Locust Street  
;; CITY: Des Moines  
;; STATE: Iowa  
;; COUNTRY: USA  
;; ZIP: 50309  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/618,911  
;; FILING DATE: Concurrently herewith  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Simon, Soma  
;; REGISTRATION NUMBER: 37,444  
;; REFERENCE/DOCKET NUMBER: 365-US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (515) 248-4800  
;; TELEFAX: (515) 248-4844

;; INFORMATION FOR SEQ ID NO: 6:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 158 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; US-08-618-911-6  
;;  
Query Match 1.8%; Score 8; DB 2; Length 158;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 436 GDDDDDD 443  
Db 55 GDDDDDD 62  
;;  
RESULT 31  
US-08-938-675A-2  
;; Sequence 2, Application US/08938675A  
;; Patent No. 6107287  
;; GENERAL INFORMATION:  
;; APPLICANT: de Lumen, Benito O.  
;; APPLICANT: Galvez, Alfredo F.  
;; TITLE OF INVENTION: Lunasin Peptides  
;; NUMBER OF SEQUENCES: 3  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
;; STREET: 75 DENISE DRIVE  
;; CITY: HILLSBOROUGH  
;; STATE: CALIFORNIA  
;; COUNTRY: USA  
;; ZIP: 94010  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/938,675A  
;; FILING DATE:  
;; CLASSIFICATION: 514  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OSMAN, RICHARD A  
;; REGISTRATION NUMBER: 36,627  
;; REFERENCE/DOCKET NUMBER: B98-003  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (650) 343-4341  
;; TELEFAX: (650) 343-4342  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 158 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-08-938-675A-2  
;;  
Query Match 1.8%; Score 8; DB 3; Length 158;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 436 GDDDDDD 443  
Db 55 GDDDDDD 62  
;;  
RESULT 32  
US-08-752-132-2  
;; Sequence 2, Application US/08752132  
;; Patent No. 5869290  
;; GENERAL INFORMATION:  
;; APPLICANT: Freeman, Kathryn

APPLICANT: Nicholas, Richard  
APPLICANT: Livi, George  
TITLE OF INVENTION: CAYAL  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Smithline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/752,132  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimm, Edward R.  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P50580  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 179 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
US-08-752-132-2

Query Match 1.8%; Score 8; DB 2; Length 179;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

436 GDDDDDD 443  
117 GDDDDDD 124

RESULT 33  
US-09-382-080-1  
Sequence 1, Application US/09382080  
Patent No. 6087333  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: DISEASE ASSOCIATED ACIDIC PROTEIN  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/382,080  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/859,937  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0308 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: TRIPINOT03  
CLONE: 2448450  
US-09-382-080-1

Query Match 1.8%; Score 8; DB 3; Length 192;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

436 GDDDDDD 443  
118 GDDDDDD 125

RESULT 34  
US-08-859-937-1  
Sequence 1, Application US/08859937  
Patent No. 6090577  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Goli, Surya K.  
TITLE OF INVENTION: DISEASE ASSOCIATED ACIDIC PROTEIN  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/859,937  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0308 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: single

TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: TRIPINOT03  
CLONE: 2448450  
US-08-859-937-1

Query Match 1.8%; Score 8; DB 3; Length 192;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GDDDDDD 443  
DB 118 GDDDDDD 125

RESULT 35  
US-08-431-080-26  
Sequence 26, Application US/08431080  
Patent No. 5698686

## GENERAL INFORMATION:

APPLICANT: Gottschling, Daniel E.  
APPLICANT: Singer, Miriam S.  
TITLE OF INVENTION: Telomerase Compositions and Methods  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TEXAS  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/431,080  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SN 08/326,781  
FILING DATE: October 20, 1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: ARCD:155/PAR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 226 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-08-431-080-26

Query Match 1.8%; Score 8; DB 1; Length 226;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
DB 142 DDDDDDDG 149

RESULT 36  
US-08-938-534-26  
Sequence 26, Application US/08938534  
Patent No. 5916752

## GENERAL INFORMATION:

APPLICANT: Gottschling, Daniel E.  
APPLICANT: Singer, Miriam S.  
TITLE OF INVENTION: Telomerase Compositions and Methods  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TEXAS  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/938,534  
FILING DATE: 26-SEP-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/431,080  
FILING DATE:  
APPLICATION NUMBER: SN 08/326,781  
FILING DATE: October 20, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: ARCD:155/PAR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 226 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-08-938-534-26

Query Match 1.8%; Score 8; DB 2; Length 226;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
DB 142 DDDDDDDG 149

RESULT 37  
US-09-345-294-26

Sequence 26, Application US/09345294  
Patent No. 6387619

## GENERAL INFORMATION:

APPLICANT: Gottschling, Daniel E.  
APPLICANT: Singer, Miriam S.  
TITLE OF INVENTION: Telomerase Compositions and Methods  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TEXAS  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/345,294  
FILING DATE: 30-JUN-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/431,080  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: ARCD-155/PAR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 226 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 26:  
-09-345-294-26  
Query Match 1.8%; Score 8; DB 4; Length 226;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 437 DDDDDDDG 444  
DB 142 DDDDDDDG 149  
RESULT 38  
US-08-469-537A-80  
Sequence 80, Application US/08469537A  
Patent No. 5843749  
GENERAL INFORMATION:  
APPLICANT: Maisongierre, et al.  
TITLE OF INVENTION: EHK AND ROR TYROSINE  
TITLE OF INVENTION: KINASES  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
STREET: 777 Old Saw Mill River Road  
CITY: Tarrytown  
STATE: NY  
COUNTRY: U.S.A.  
ZIP: 10591  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,537A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/406,247  
FILING DATE: 17-MAR-1995  
APPLICATION NUMBER: USSN 08/144,992  
FILING DATE: 28-OCT-1993  
APPLICATION NUMBER: USSN 07/736,559  
FILING DATE: 26-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Kempster, Ph.D., Gail M  
REGISTRATION NUMBER: 32,143  
REFERENCE/DOCKET NUMBER: REG 070C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 914-345-7400  
TELEFAX: 914-345-7721  
TELEX:  
INFORMATION FOR SEQ ID NO: 80:

SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-469-537A-80  
Query Match 1.8%; Score 8; DB 2; Length 251;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 191 SYGVVLM 198  
DB 109 SYGVVLM 116  
RESULT 39  
US-08-469-537A-81  
Sequence 81, Application US/08469537A  
Patent No. 5843749  
GENERAL INFORMATION:  
APPLICANT: Maisongierre, et al.  
TITLE OF INVENTION: EHK AND ROR TYROSINE  
TITLE OF INVENTION: KINASES  
NUMBER OF SEQUENCES: 107  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
STREET: 777 Old Saw Mill River Road  
CITY: Tarrytown  
STATE: NY  
COUNTRY: U.S.A.  
ZIP: 10591  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,537A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/406,247  
FILING DATE: 17-MAR-1995  
APPLICATION NUMBER: USSN 08/144,992  
FILING DATE: 28-OCT-1993  
APPLICATION NUMBER: USSN 07/736,559  
FILING DATE: 26-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Kempster, Ph.D., Gail M  
REGISTRATION NUMBER: 32,143  
REFERENCE/DOCKET NUMBER: REG 070C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 914-345-7400  
TELEFAX: 914-345-7721  
TELEX:  
INFORMATION FOR SEQ ID NO: 81:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-469-537A-81  
Query Match 1.8%; Score 8; DB 2; Length 251;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 191 SYGVVLM 198  
DB 109 SYGVVLM 116

```
RESULT 40
US-08-852-743-5
; Sequence 5, Application US/08852743
; Patent No. 5830699
; GENERAL INFORMATION:
; APPLICANT: Force, Thomas
; APPLICANT: Kyriakis, John M.
; APPLICANT: Pombo, Celia M.
; APPLICANT: Bonventre, Joseph
; TITLE OF INVENTION: SOK-1 AND METHODS OF USE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,743
; FILING DATE: 7-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/016,774
; FILING DATE: 7-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00786/327001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-852-743-5

Query Match 1.8%; Score 8; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
DB 165 WMAPEVIQ 172

RESULT 41
US-09-185-370-5
; Sequence 5, Application US/09185370
; Patent No. 6093560
; GENERAL INFORMATION:
; APPLICANT: Force, Thomas
; APPLICANT: Kyriakis, John M.
; APPLICANT: Pombo, Celia M.
; APPLICANT: Bonventre, Joseph
; TITLE OF INVENTION: SOK-1 AND METHODS OF USE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
```

```
ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/185,370
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/852,743
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00786/327001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-185-370-5
```

```
Query Match 1.8%; Score 8; DB 3; Length 270;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIQ 179
DB 165 WMAPEVIQ 172
```

```
RESULT 42
US-09-485-077A-17
; Sequence 17, Application US/09485077A
; Patent No. 6458590
; GENERAL INFORMATION:
; APPLICANT: Mukherjee, Anil
; APPLICANT: Kundu, Gopal
; APPLICANT: Panda, Dibyendu
; TITLE OF INVENTION: Methods and Compositions for Treatment of Restenosis
; FILE REFERENCE: NIH-05047
; CURRENT APPLICATION NUMBER: US/09/485,077A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: PCT/US98/16569
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: 60/054,967
; PRIOR FILING DATE: 1997-07-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Rattus No. 6458590vegicus
; US-09-485-077A-17
```

```
Query Match 1.8%; Score 8; DB 4; Length 317;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444
DB 88 DDDDDDDG 95
```

```
RESULT 43
US-09-211-930-3
```

```
; Sequence 3, Application US/09211930
; Patent No. 5962265
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 5962265r1s
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION
; TITLE OF INVENTION: SERINE/THREONINE KINASE
; FILE REFERENCE: PHM.70296
; CURRENT APPLICATION NUMBER: US/09/211,930
; CURRENT FILING DATE: 1998-12-15
; EARLIER APPLICATION NUMBER: GB 9726851.0
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-211-930-3

Query Match      1.8%; Score 8; DB 2; Length 416;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 172 WMAPEVIO 179
Db 185 WMAPEVIO 192

RESULT 44
; US-09-211-930-11
; Sequence 11, Application US/09211930
; Patent No. 5962265
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 5962265r1s
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION
; TITLE OF INVENTION: SERINE/THREONINE KINASE
; FILE REFERENCE: PHM.70296
; CURRENT APPLICATION NUMBER: US/09/211,930
; CURRENT FILING DATE: 1998-12-15
; EARLIER APPLICATION NUMBER: GB 9726851.0
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-211-930-11

Query Match      1.8%; Score 8; DB 2; Length 416;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 172 WMAPEVIO 179
Db 185 WMAPEVIO 192

RESULT 45
; US-09-340-993-3
; Sequence 3, Application US/09340993
; Patent No. 6034228
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 6034228r1s
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION
; TITLE OF INVENTION: SERINE/THREONINE KINASE
; FILE REFERENCE: PHM.70296.N1
; CURRENT APPLICATION NUMBER: US/09/340,993
```

```
; CURRENT FILING DATE: 1999-06-25
; EARLIER APPLICATION NUMBER: GB 9726851.0 & US 09/211,930
; EARLIER FILING DATE: 1997-12-19 & 1998-12-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-340-993-3

Query Match      1.8%; Score 8; DB 3; Length 416;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 172 WMAPEVIO 179
Db 185 WMAPEVIO 192

RESULT 46
; US-09-340-993-11
; Sequence 11, Application US/09340993
; Patent No. 6034228
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 6034228r1s
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION
; TITLE OF INVENTION: SERINE/THREONINE KINASE
; FILE REFERENCE: PHM.70296.N1
; CURRENT APPLICATION NUMBER: US/09/340,993
; CURRENT FILING DATE: 1999-06-25
; EARLIER APPLICATION NUMBER: GB 9726851.0 & US 09/211,930
; EARLIER FILING DATE: 1997-12-19 & 1998-12-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-340-993-11

Query Match      1.8%; Score 8; DB 3; Length 416;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 172 WMAPEVIO 179
Db 185 WMAPEVIO 192

RESULT 47
; US-09-468-442-3
; Sequence 3, Application US/09468442
; Patent No. 6300098
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 6300098r1s
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION
; TITLE OF INVENTION: SERINE/THREONINE KINASE
; FILE REFERENCE: PHM.70296.N1
; CURRENT APPLICATION NUMBER: US/09/468,442
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: US 09/340,993
; EARLIER FILING DATE: 1999-06-25
; EARLIER APPLICATION NUMBER: GB 9726851.0
; EARLIER FILING DATE: 1997-12-19
; EARLIER APPLICATION NUMBER: US 09/211,930
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 416
```

TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-468-442-3

Query Match 1.8%; Score 8; DB 4; Length 416;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 172 WMAPEVIO 179  
Db 185 WMAPEVIO 192

RESULT 48  
US-09-468-442-11

Sequence 11, Application US/09468442  
Patent No. 6300098  
GENERAL INFORMATION:  
APPLICANT: Tyrell E. No. 6300098-18  
APPLICANT: William Craig Moore  
TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION-SERINE/THREONINE KINASE  
FILE REFERENCE: PHM.70296.NI  
CURRENT APPLICATION NUMBER: US/09/468,442  
EARLIER FILING DATE: 1999-12-21  
EARLIER APPLICATION NUMBER: US 09/340,993  
EARLIER FILING DATE: 1999-06-25  
EARLIER APPLICATION NUMBER: GB 9726851.0  
EARLIER FILING DATE: 1997-12-19  
EARLIER APPLICATION NUMBER: US 09/211,930  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 11  
LENGTH: 416  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-468-442-11

Query Match 1.8%; Score 8; DB 4; Length 416;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 172 WMAPEVIO 179  
Db 185 WMAPEVIO 192

RESULT 49  
US-08-712-709-8

Sequence 8, Application US/08712709  
Patent No. 5863780  
GENERAL INFORMATION:  
APPLICANT: Au-Young, Janice  
APPLICANT: Guegler, Karl J.  
APPLICANT: Hawkins, Phillip R.  
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/712,709  
FILING DATE: Filed Herewith

ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PP-0118 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-845-4166  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 487 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1117791  
US-08-712-709-8

Query Match 1.8%; Score 8; DB 2; Length 487;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 172 WMAPEVIO 179  
Db 190 WMAPEVIO 197

RESULT 50  
US-09-111-444-8

Sequence 8, Application US/09111444  
Patent No. 6045792  
GENERAL INFORMATION:  
APPLICANT: Au-Young, Janice  
APPLICANT: Guegler, Karl J.  
APPLICANT: Hawkins, Phillip R.  
TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: U.S.  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/111,444  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/712,709  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PP-0118 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-845-4166  
TELEFAX: 415-845-0555  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 487 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1117791

US-09-111-444-8

Query Match 1.8%; Score 8; DB 3; Length 487;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIO 179  
Db 190 WMAPEVIO 197

RESULT 51

US-09-541-228-8  
; Sequence 8; Application US/09541228  
; Patent No. 6232077  
; GENERAL INFORMATION:  
; APPLICANT: Au-Young, Janice  
; APPLICANT: Gugler, Karl J.  
; APPLICANT: Hawkins, Phillip R.  
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: U.S.  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/541,228  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/712,709  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0118 'US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 487 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; IMMEDIATE SOURCE:  
; LIBRARY: Genbank  
; CLONE: 1117791  
; US-09-541-228-8

Query Match 1.8%; Score 8; DB 4; Length 487;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 WMAPEVIO 179  
Db 190 WMAPEVIO 197

RESULT 52

US-09-529-279-4  
; Sequence 4; Application US/09529279  
; Patent No. 6451617  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO

APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/09/529,279  
; CURRENT FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-279-4

Query Match 1.8%; Score 8; DB 4; Length 579;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153  
Db 170 VLKICDFG 177

RESULT 53

US-09-529-279-15  
; Sequence 15; Application US/09529279  
; Patent No. 6451617  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/09/529,279  
; CURRENT FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 590  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-279-15

Query Match 1.8%; Score 8; DB 4; Length 590;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 VLKICDFG 153  
Db 170 VLKICDFG 177

RESULT 54

US-08-469-537A-107  
; Sequence 107; Application US/08469537A  
; Patent No. 5843749  
; GENERAL INFORMATION:  
; APPLICANT: Maisongier, et al.  
; TITLE OF INVENTION: EHK AND ROR TYROSINE  
; NUMBER OF SEQUENCES: 107  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
; STREET: 777 Old Saw Mill River Road  
; CITY: Tarrytown



STATE: NY  
COUNTRY: U.S.A.  
ZIP: 10591  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,537A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/406,247  
FILING DATE: 17-MAR-1995  
APPLICATION NUMBER: USSN 08/144,992  
FILING DATE: 28-OCT-1993  
APPLICATION NUMBER: USSN 07/736,559  
FILING DATE: 26-JUL-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Kempier, Ph.D., Gall M  
REGISTRATION NUMBER: 32,143  
REFERENCE/DOCKET NUMBER: REG 070C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 914-345-7400  
TELEFAX: 914-345-7721  
TELEX:  
INFORMATION FOR SEQ ID NO: 107:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 943 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
FEATURE:  
NAME/KEY: Human ROR2  
LOCATION: 1--943  
OTHER INFORMATION:  
US-08-469-537A-107

Query Match 1.8%; Score 8; DB 2; Length 943;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 191 SYGVVLWE 198  
|||||  
677 SYGVVLWE 684

RESULT 55  
US-08-431-080-28  
Sequence 28, Application US/08431080  
Patent No. 5698686  
GENERAL INFORMATION:  
APPLICANT: Gottschling, Daniel E.  
APPLICANT: Singer, Miriam S.  
TITLE OF INVENTION: Telomerase Compositions and Methods  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TEXAS  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/431,080

FILING DATE: Concurrently Herewith  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SN 08/326,781  
FILING DATE: October 20, 1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: ARCD:155/PAR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1085 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-431-080-28

Query Match 1.8%; Score 8; DB 1; Length 1085;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 437 DDDDDDDG 444  
|||||  
Db 142 DDDDDDDG 149

RESULT 56  
US-08-938-534-28  
Sequence 29, Application US/08938534  
Patent No. 5916752  
GENERAL INFORMATION:  
APPLICANT: Gottschling, Daniel E.  
APPLICANT: Singer, Miriam S.  
TITLE OF INVENTION: Telomerase Compositions and Methods  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: TEXAS  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/938,534  
FILING DATE: 26-SEP-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/431,080  
FILING DATE:  
APPLICATION NUMBER: SN 08/326,781  
FILING DATE: October 20, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165  
REFERENCE/DOCKET NUMBER: ARCD:155/PAR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679  
TELEX: 79-0924  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1085 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-938-534-28

Query Match 1.8%; Score 8; DB 2; Length 1085;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
Db 142 DDDDDDDG 149

RESULT 57  
US-09-345-294-28

Sequence 28, Application US/09345294  
Patent No. 6387619

GENERAL INFORMATION:

APPLICANT: Gotschling, Daniel E.  
Singer, Miriam S.

TITLE OF INVENTION: Telomerase Compositions and Methods  
NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433

CITY: Houston  
STATE: TEXAS

COUNTRY: UNITED STATES OF AMERICA  
ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII

SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/345,294  
FILING DATE: 30-Jun-1999

CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/431,080  
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Parker, David L.  
REGISTRATION NUMBER: 32,165

REFERENCE/DOCKET NUMBER: ARCD:155/PAR  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000  
TELEFAX: (713) 789-2679

TELEX: 79-0924

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 1085 amino acids  
\* TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 28:

Query Match 1.8%; Score 8; DB 4; Length 1085;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDDG 444  
Db 142 DDDDDDDG 149

RESULT 58  
US-08-482-228-215

Sequence 215, Application US/08482228  
Patent No. 5968753

GENERAL INFORMATION:

APPLICANT: Tseng-Law, Janet

APPLICANT: Kobori, Joan A.  
APPLICANT: Al-Abdaly, Fahad A.

APPLICANT: Guillermo, Roy L.  
APPLICANT: Helgeson, Sam L.

APPLICANT: Deans, Robert J.  
TITLE OF INVENTION: POSITIVE AND POSITIVE/NEGATIVE CELL  
TITLE OF INVENTION: SELECTION MEDIATED BY PEPTIDE RELEASE

NUMBER OF SEQUENCES: 215  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Janice Guthrie, Ph.D.  
STREET: P.O. Box 15210

CITY: Irvine  
STATE: California

COUNTRY: USA  
ZIP: 92713-5210

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/482,228  
FILING DATE: 07-JUN-1995

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:

NAME: Guthrie, Janice  
REGISTRATION NUMBER: 35,170

REFERENCE/DOCKET NUMBER: IT-4630CIP3  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (714) 440-5353  
TELEFAX: (714) 553-1952

INFORMATION FOR SEQ ID NO: 215:

SEQUENCE CHARACTERISTICS:

LENGTH: 11 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: peptide

US-08-482-228-215

Query Match 1.5%; Score 7; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
Db 1 DDDDDDD 7

RESULT 59  
US-08-482-528-215

Sequence 215, Application US/08482528  
Patent No. 6017719

GENERAL INFORMATION:

APPLICANT: Tseng-Law, Janet  
APPLICANT: Kobori, Joan A.

APPLICANT: Al-Abdaly, Fahad A.  
APPLICANT: Guillermo, Roy L.

APPLICANT: Helgeson, Sam L.  
APPLICANT: Deans, Robert J.

TITLE OF INVENTION: POSITIVE AND POSITIVE/NEGATIVE CELL  
TITLE OF INVENTION: SELECTION MEDIATED BY PEPTIDE RELEASE  
NUMBER OF SEQUENCES: 215

CORRESPONDENCE ADDRESS:

ADDRESSEE: Janice Guthrie, Ph.D.  
STREET: P.O. Box 15210

CITY: Irvine  
STATE: California

COUNTRY: USA  
ZIP: 92713-5210

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,528  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Guthrie, Janice  
REGISTRATION NUMBER: 35,170  
REFERENCE/DOCKET NUMBER: IT-4630CIP4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (714) 440-5353  
TELEFAX: (714) 553-1952  
INFORMATION FOR SEQ ID NO: 215:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-482-528-215

Query Match 1.5%; Score 7; DB 3; Length 11;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDD 443  
Db 1 DDDDDD 7

RESULT 60  
US-08-199-776-23  
Sequence 23, Application US/08199776  
Patent No. 5594120  
GENERAL INFORMATION:  
APPLICANT: Brenner, Michael B.  
APPLICANT: Parker, Christina M.  
TITLE OF INVENTION: No. 5594120e1 Integrin alpha subunit  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield and Sacks, P.C.  
STREET: 600 Atlantic Avenue  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/199,776  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Plumer, Elizabeth R.  
REGISTRATION NUMBER: 36,637  
REFERENCE/DOCKET NUMBER: B0801/7020  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-720-3500  
TELEFAX: 617-720-2441  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: synthetic peptide  
US-08-199-776-23

Query Match 1.5%; Score 7; DB 1; Length 12;  
Best Local Similarity 100.0%; Pred. No. 9.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDD 443  
Db 3 DDDDDD 9

RESULT 61  
US-08-663-731-23  
Sequence 23, Application US/08663731  
Patent No. 6057423  
GENERAL INFORMATION:  
APPLICANT: Brenner, Michael B.  
APPLICANT: Parker, Christina M.  
TITLE OF INVENTION: No. 6057423e1 Integrin alpha subunit  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield and Sacks, P.C.  
STREET: 600 Atlantic Avenue  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/663,731  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/199,776  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Plumer, Elizabeth R.  
REGISTRATION NUMBER: 36,637  
REFERENCE/DOCKET NUMBER: B0801/7020  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-720-3500  
TELEFAX: 617-720-2441  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: YES  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: synthetic peptide  
US-08-663-731-23

Query Match 1.5%; Score 7; DB 3; Length 12;  
Best Local Similarity 100.0%; Pred. No. 9.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDD 443  
Db 3 DDDDDD 9

RESULT 62  
US-08-879-338-23  
Sequence 23, Application US/08879338A  
Patent No. 6063906  
GENERAL INFORMATION:  
APPLICANT: Brenner, Michael B.

```

; APPLICANT: Parker, Christina M.
; TITLE OF INVENTION: Antibodies to No. 6063906e1 Integrin Alpha
; FILE REFERENCE: B0801/7080/ERP
; CURRENT FILING DATE: 1997-06-20
; EARLIER APPLICATION NUMBER: US 08/663,731
; EARLIER FILING DATE: 1996-06-14
; EARLIER APPLICATION NUMBER: US 08/199,776
; EARLIER FILING DATE: 1994-02-18
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
; US-08-879-338-23

Query Match
Best Local Similarity 1.5%; Score 7; DB 3; Length 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
Db 3 DDDDDDD 9

RESULT 63
US-09-293-238B-23
; Sequence 23, Application US/09293238B
; Patent No. 6455042
; GENERAL INFORMATION:
; APPLICANT: Brenner, Michael B.
; TITLE OF INVENTION: A Method of Treating Ulcerative Colitis
; FILE REFERENCE: 10560/7005/ERP
; CURRENT FILING DATE: 1999-04-16
; CURRENT APPLICATION NUMBER: US/09/293,238B
; PRIOR FILING DATE: 1997-06-20
; PRIOR APPLICATION NUMBER: US 08/663,731
; PRIOR FILING DATE: 1996-06-14
; PRIOR APPLICATION NUMBER: US 08/199,776
; PRIOR FILING DATE: 1994-02-18
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
; US-09-293-238B-23

Query Match
Best Local Similarity 1.5%; Score 7; DB 4; Length 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
Db 3 DDDDDDD 9

RESULT 64
PCT-US95-02044-23
; Sequence 23, Application PC/TUS9502044
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Novel Integrin alpha subunit
```

```

; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Wolf, Greenfield and Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02044
; FILING DATE: herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/199,776
; FILING DATE: 18 February 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-2441
; TELEFAX: 617-720-3500
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: synthetic peptide
; PCT-US95-02044-23

Query Match
Best Local Similarity 1.5%; Score 7; DB 5; Length 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443
Db 3 DDDDDDD 9

RESULT 65
US-09-625-570-4
; Sequence 4, Application US/09625570
; Patent No. 6437095
; GENERAL INFORMATION:
; APPLICANT: Lillie, Hauke
; APPLICANT: Richter, Susanne
; APPLICANT: Rudolph, Rainer
; APPLICANT: Stubentrauch, Kay-Gunnar
; TITLE OF INVENTION: Chimeric polypeptides, method for production and uses
; FILE REFERENCE: Case 20393
; CURRENT FILING DATE: 2000-07-26
; CURRENT APPLICATION NUMBER: US/09/625,570
; PRIOR FILING DATE: 1999-08-02
; PRIOR APPLICATION NUMBER: EP99115022.8
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 4
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide D10SCP
```

US-09-625-570-4

Query Match 1.5%; Score 7; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDDD 443  
DB 1 DDDDDDD 7

RESULT 66

US-09-039-780A-80

Sequence 80, Application US/09039780A

Patent No. 6376248

GENERAL INFORMATION:

APPLICANT: HAMLEY-NELSON, PAMELA

LAN, JIANQING

SHIH, POJEN

JESSE, JOEL A.

SCHIFFERLI, KEVIN P.

TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS

NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: GREENLEE, WINNER &amp; SULLIVAN

STREET: 5370 MANHATTAN CIRCLE, SUITE 201

CITY: BOULDER

STATE: CO.

COUNTRY: US

ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/039,780A

FILING DATE: 16-Mar-1998

CLASSIFICATION: &lt;Unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: SULLIVAN, SALLY A.

REGISTRATION NUMBER: 32,064

REFERENCE/DOCKET NUMBER: 32-95C

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303)499-8080

TELEFAX: (303)499-8089

INFORMATION FOR SEQ ID NO: 80:

SEQUENCE CHARACTERISTICS:

LENGTH: 14 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: Modified-site

LOCATION: 14

OTHER INFORMATION: /product= "OTHER"

/note= "G AT POSITION 14 CAN BE ABSENT"

US-09-039-780A-80

Query Match 1.5%; Score 7; DB 4; Length 14;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 438 DDDDDG 444  
DB 8 DDDDDG 14

RESULT 67

US-09-039-780A-87

Sequence 87, Application US/09039780A

Patent No. 6376248

GENERAL INFORMATION:

APPLICANT: HAMLEY-NELSON, PAMELA

LAN, JIANQING

SHIH, POJEN

JESSE, JOEL A.

SCHIFFERLI, KEVIN P.

TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS

NUMBER OF SEQUENCES: 120

CORRESPONDENCE ADDRESS:

ADDRESSEE: GREENLEE, WINNER &amp; SULLIVAN

STREET: 5370 MANHATTAN CIRCLE, SUITE 201

CITY: BOULDER

STATE: CO

COUNTRY: US

ZIP: 80303

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/039,780A

FILING DATE: 16-Mar-1998

CLASSIFICATION: &lt;Unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: SULLIVAN, SALLY A.

REGISTRATION NUMBER: 32,064

REFERENCE/DOCKET NUMBER: 32-95C

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303)499-8080

TELEFAX: (303)499-8089

INFORMATION FOR SEQ ID NO: 87:

SEQUENCE CHARACTERISTICS:

LENGTH: 30 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: Modified-site

LOCATION: 30

OTHER INFORMATION: /product= "OTHER"

/note= "G AT POSITION 30 CAN BE ABSENT"

US-09-039-780A-87

Query Match 1.5%; Score 7; DB 4; Length 30;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 438 DDDDDG 444  
DB 24 DDDDDG 30

RESULT 68

US-09-376-113-3

Sequence 3, Application US/09376113

Patent No. 6451992

GENERAL INFORMATION:

APPLICANT: Cupp, Eddie Wayne

Cupp, Mary Smith

TITLE OF INVENTION: Antithrombin Nucleotides and Proteins

TITLE OF INVENTION: from Horn Fly

FILE REFERENCE: 5721-10

CURRENT APPLICATION NUMBER: US/09/376,113

CURRENT FILING DATE: 1999-08-17  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 30  
TYPE: PRT  
ORGANISM: Haemotobia irritans  
US-09-376-113-3

Query Match 1.5%; Score 7; DB 4; Length 30;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443  
DB 10 DDDDDD 16

RESULT 69  
US-08-290-736C-9  
Sequence 9, Application US/08290736C  
Patent No. 6294174

GENERAL INFORMATION:

APPLICANT: KRSMANOVIC, VELLIBOR  
COSIC, IRENA  
BIOUARD, JEAN-MICHEL  
HEARN, MILTON TM

TITLE OF INVENTION: PEPTIDES IMMUNOLOGICALLY RELATED TO  
PROTEINS OF A VIRAL AGENT AND THEIR BIOLOGICAL APPLICATION

NUMBER OF SEQUENCES: 48

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHAYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VA

COUNTRY: USA  
ZIP: 22201

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/290 736C

FILING DATE: 16-NO. 6294174-1994

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/FR93/00171

FILING DATE: 19-FEB-1993

APPLICATION NUMBER: FR92/01883

FILING DATE: 19-FEB-1992

ATTORNEY/AGENT INFORMATION:

NAME: SADOFF, B.J.

REGISTRATION NUMBER: 36663

REFERENCE/DOCKET NUMBER: 1721-3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 7038164000

TELEFAX: 7038164100

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 50 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Peptide

LOCATION: 1..50

SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Query Match 1.5%; Score 7; DB 4; Length 50;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443  
DB 4 DDDDDD 10

RESULT 70  
US-08-290-736C-8  
Sequence 8, Application US/08290736C  
Patent No. 6294174

GENERAL INFORMATION:

APPLICANT: KRSMANOVIC, VELLIBOR  
COSIC, IRENA  
BIOUARD, JEAN-MICHEL  
HEARN, MILTON TM

TITLE OF INVENTION: PEPTIDES IMMUNOLOGICALLY RELATED TO  
PROTEINS OF A VIRAL AGENT AND THEIR BIOLOGICAL APPLICATION

NUMBER OF SEQUENCES: 48

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHAYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VA

COUNTRY: USA  
ZIP: 22201

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/290 736C

FILING DATE: 16-NO. 6294174-1994

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/FR93/00171

FILING DATE: 19-FEB-1993

APPLICATION NUMBER: FR92/01883

FILING DATE: 19-FEB-1992

ATTORNEY/AGENT INFORMATION:

NAME: SADOFF, B.J.

REGISTRATION NUMBER: 36663

REFERENCE/DOCKET NUMBER: 1721-3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 7038164000

TELEFAX: 7038164100

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 73 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Peptide

LOCATION: 1..73

SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Query Match 1.5%; Score 7; DB 4; Length 73;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 DDDDDD 443  
DB 40 DDDDDD 46

RESULT 71  
US-08-117-952-757  
Sequence 757, Application US/08117952  
Patent No. 5851760  
GENERAL INFORMATION:  
APPLICANT: Evans, Glen A.

APPLICANT: Smith, Michael W.  
TITLE OF INVENTION: METHOD FOR GENERATION OF SEQUENCE  
NUMBER OF SEQUENCES: 797  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
STREET: 444 South Flower Street, Suite 2000  
CITY: Los Angeles  
STATE: CA  
COUNTRY: USA  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/117,952  
FILING DATE: 07-SEP-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/078,471  
FILING DATE: 15-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Reiter, Stephen E.  
REGISTRATION NUMBER: 31,192  
REFERENCE/DOCKET NUMBER: P41 9423  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-546-4737  
TELEFAX: 619-546-9392  
INFORMATION FOR SEQ ID NO: 757:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 76 amino acids  
TYPE: amino acid  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-08-117-952-757

Query Match 1.5%; Score 7; DB 2; Length 76;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 129 VHRDLK 135  
Db 59 VHRDLK 65

RESULT 72  
US-09-376-113-2  
Sequence 2, Application US/09376113  
Patent No. 6451992  
GENERAL INFORMATION:  
APPLICANT: Cupp, Eddie Wayne  
APPLICANT: Cupp, Mary Smith  
TITLE OF INVENTION: Antichrombin Nucleotides and Proteins  
FILE REFERENCE: 5721-10  
CURRENT APPLICATION NUMBER: US/09/376,113  
CURRENT FILING DATE: 1999-08-17  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO: 2  
LENGTH: 81  
TYPE: PRT  
ORGANISM: Haematobia irritans  
US-09-376-113-2

Query Match 1.5%; Score 7; DB 4; Length 81;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 437 DDDDDDD 443

Db 10 DDDDDDD 16

RESULT 73  
US-08-370-225-20  
Sequence 20, Application US/08370225  
Patent No. 5580736  
GENERAL INFORMATION:  
APPLICANT: Brent, Roger  
APPLICANT: Gyuris, Jeno  
APPLICANT: Golemis, Erica  
TITLE OF INVENTION: Interaction Trap System for Isolating  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 555X  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: Wordperfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/370,225  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/969,038  
FILING DATE: 10/30/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/143001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-370-225-20

Query Match 1.5%; Score 7; DB 1; Length 86;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 129 VHRDLK 135  
Db 47 VHRDLK 53

RESULT 74  
US-08-461-859-20  
Sequence 20, Application US/08461859  
Patent No. 5786169  
GENERAL INFORMATION:  
APPLICANT: Brent, Roger  
APPLICANT: Gyuris, Jeno  
APPLICANT: Golemis, Erica  
TITLE OF INVENTION: Interaction Trap System for Isolating  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street

Sun May 4 10:01:03 2003

us-09-757-982-5.oligo.ra1

Page 28

CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 50Z or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,859  
FILING DATE: June 5, 1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/370,225  
FILING DATE: January 9, 1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/969,038  
FILING DATE: October 30, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lech, Karen F.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/143002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-461-859-20

Query Match 1.5%; Score 7; DB 1; Length 86;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VIHRLDK 135  
DB 47 VIHRLDK 53

RESULT 75  
PCT-US93-10069-20  
Sequence 20, Application PC/TUS9310069  
GENERAL INFORMATION:  
APPLICANT: Brent, Roger  
APPLICANT: Gyuris, Jeno  
APPLICANT: Golemis, Erica  
TITLE OF INVENTION: Interaction Trap System for Isolating  
TITLE OF INVENTION: Novel Proteins  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 50Z or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/10069  
FILING DATE: 20-OCT-1993  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/969,038  
FILING DATE: 10/30/92

ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/143001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 86  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
PCT-US93-10069-20

Query Match 1.5%; Score 7; DB 5; Length 86;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 VIHRLDK 135  
DB 47 VIHRLDK 53

Search completed: May 1, 2003, 20:56:36  
Job time: 31 secs